

Arthur Faria Porto

Estudo de Técnicas de Registro e Reconstrução de Imagens Aplicadas ao Problema de Super-Resolução

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Inteligência Computacional.

Orientador: André Paim Lemos

Belo Horizonte, MG
2015

Resumo

Este trabalho apresenta uma avaliação experimental de diferentes técnicas de registro e reconstrução de imagens aplicadas ao problema de Super-Resolução. O problema da Super-Resolução consiste em gerar uma imagem de alta resolução a partir de um conjunto de imagens de baixa resolução de uma mesma cena. A abordagem de Super-Resolução pode ser dividida em duas etapas: registro e reconstrução. Após registrar as imagens de baixa resolução sobre um mesmo plano de alta resolução, a fase de reconstrução combina essas imagens formando uma única imagem de alta resolução. Quatro métodos de registro (*Keren*, *Marcel*, *Lucchese* e *Vandewalle*) e cinco métodos de reconstrução (*IBP*, *TV*, *RSR*, *NC* and *G-PMSR*) de imagens foram avaliados. O objetivo do trabalho é avaliar, dentre os métodos considerados, qual é a combinação de técnicas capaz de gerar uma imagem de alta resolução que melhor representa a cena observada. Inicialmente foram realizados experimentos em um ambiente controlado, gerando-se artificialmente imagens de baixa resolução, para avaliar o desempenho das técnicas de registro e reconstrução consideradas. Após a seleção dos métodos que se destacaram no experimento inicial, um novo experimento, com imagens de baixa resolução reais, foi realizado para validar a combinação de métodos escolhida. Os resultados obtidos demonstram a viabilidade da utilização da combinação dos métodos *Keren* e *G-PMSR* em problemas de Super-Resolução. As imagens de alta resolução geradas foram capazes de recuperar detalhes da cena que não eram perceptíveis apenas com as imagens de baixa resolução.

Palavras-chave: Super-Resolução, Registro de Imagens, Reconstrução de Imagens.

Abstract

This work presents an experimental evaluation of different image registration and reconstruction techniques applied to the Super-Resolution problem. The Super-Resolution problem aims to generate a high resolution image from a set of low-resolution images of the same scene. The Super-Resolution approach can be divided into two phases: registration and reconstruction. After registering the low resolution images on a single high resolution grid, the reconstruction phase combines these images into a single image of high resolution. Four image registration methods (*Keren*, *Marcel*, *Lucchese* and *Vandewalle*) and five image reconstruction methods were evaluated (*IBP*, *TV*, *RSR*, *NC* and *G-PMSR*). The objective is to assess, among the considered methods, which is the best combination of techniques able to generate a high resolution image with a better expression of the observed scene. Initially experiments were conducted in a controlled environment, generating artificially low resolution images, to evaluate the performance of registration and reconstruction techniques considered. After the selection of the excelled methods in the initial experiment, a new experiment, with real low resolution images was performed to validate the combination of chosen methods. The results demonstrate the feasibility of using the combination of the *Keren* and *G-PMSR* methods in Super-Resolution problems. The high resolution images generated were able to retrieve details of the scene that were not perceptible with the low resolution images alone.

Keywords: Super-Resolution, Image Registration, Image Reconstruction.

Agradecimentos

A Jesus, o Cristo, pois Ele é antes de todas as coisas, e nele tudo subsiste.

A minha esposa Ana Carolina, por amar quem eu sou e estar sempre ao meu lado.

Aos meus pais, Tarcísio e Jurema, por me ensinarem no caminho e me ajudarem nas minhas decisões.

Ao André, por ser o melhor irmão que eu preciso e me instigar a pensar.

A minha avó Raymunda, pelas orações e acolhimento.

Ao meu orientador, Prof. André Paim Lemos, pela sempre oportuna disposição em me auxiliar e trazer direcionamento ao trabalho.

Aos demais professores do PPGEE, pelo compartilhar do conhecimento.

Aos funcionários do PPGEE, pela colaboração e atenção.

Ao CNPq, pelo apoio financeiro, N^o 134288/2014-9.

Ao Instituto Federal do Norte de Minas Gerais, especialmente ao campus Salinas pelo apoio e suporte a realização do trabalho.

Aos companheiros de tripulação da Sétima, pelo companheirismo.

Aos meus colegas de mestrado e em especial os membros do D!FCOM, pelas discussões e convivência.

Aos demais familiares e amigos, pelo apoio e torcida.

Sumário

Sumário	v
Lista de Figuras	vi
Lista de Tabelas	viii
Lista de Abreviaturas e Siglas	x
1 Introdução	1
1.1 Motivação e Relevância	1
1.2 Definição do Problema de Super-Resolução	4
1.3 Abordagem	5
1.4 Objetivos	6
1.4.1 Objetivos Gerais	6
1.4.2 Objetivos Específicos	6
1.5 Contribuição	6
1.6 Organização do Texto	7
2 Revisão de Literatura	9
2.1 Introdução	9
2.2 Métodos de Registro	12
2.2.1 Keren	12
2.2.2 Marcel	14
2.2.3 Lucchese	15
2.2.4 Vandewalle	16
2.3 Métodos de Reconstrução	18
2.3.1 Iterative Back Projection - IBP	18
2.3.2 Total Variation - TV	19
2.3.3 Robust Super-Resolution - RSR	20
2.3.4 Normalized Convolution - NC	21
2.3.5 G-PMSR	22
3 Experimentos	25
3.1 Introdução	25
3.2 Experimentos com imagens simuladas	26

3.2.1	Metodologia	26
3.2.2	Métricas de avaliação	26
3.2.3	Dados	28
3.2.4	Casos de Teste	29
3.2.5	Resultados	32
3.3	Experimentos com imagens reais	48
3.3.1	Metodologia	48
3.3.2	Dados	48
3.3.3	Resultados	49
3.4	Conclusão	52
4	Conclusão	57
4.1	Trabalhos Futuros	58
	Referências bibliográficas	59

Lista de Figuras

1.1	Alvo de teste USAF 1951 usado para determinar a resolução espacial de sensores e sistemas de imagens. Imagens com dimensões correspondentes, no entanto, (a) representa uma imagem com baixa resolução espacial, enquanto (b) possui uma resolução quatro vezes maior que (a).	2
1.2	Ilustração básica da reconstrução de uma imagem em alta resolução através de múltiplas imagens de BR apresentando deslocamentos a nível de sub-píxel. . . .	3
1.3	Modelo de observação de um sistema de captura de imagens relacionando uma cena em alta resolução e as imagens de BR geradas pela câmera (adaptado de Milanfar (2010)).	4
2.1	Interpolação de múltiplas imagens de BR, com deslocamentos de sub-píxel, em um imagem com resolução aumentada. (a) conjunto de imagens de BR; (b) as imagens organizadas de acordo com os seus deslocamentos; (c) combinação das imagens em um plano de resolução aumentada; (d) resultado do processo de interpolação.	10
2.2	Interpolação de múltiplas imagens de BR, com deslocamentos inteiros e de sub-píxel, em um imagem com resolução aumentada. (a) conjunto de imagens de BR; (b) as imagens organizadas de acordo com os seus deslocamentos; (c) combinação das imagens em um plano de resolução aumentada; (d) resultado do processo de interpolação.	10
2.3	Interpolação de múltiplas imagens de BR, com deslocamentos de sub-píxel, em um imagem com resolução aumentada, com a quantidade de píxeis do conjunto de BR igual a quantidade de píxeis da imagem de AR. (a) conjunto de imagens de BR; (b) as imagens organizadas de acordo com os seus deslocamentos; (c) combinação das imagens em um plano de resolução aumentada; (d) resultado do processo de interpolação.	11
2.4	Estimativa de Rotação. (a) Valores de frequência da imagem de referência para $0.1\rho < r < \rho$. (b) Valores de frequência da imagem rotacionada para $0.1\rho < r < \rho$. (c) Valor médio em função do ângulo $h(\alpha)$ para $ G_0(\mathbf{u}) $ e $ G_1(\mathbf{u}) $. Fonte: Vandewalle (2006).	17
2.5	Diagrama semântico do algoritmo de super resolução do IBP (Adaptado de Park (2003)).	18
3.1	Exemplos de imagens utilizadas na primeira fase se experimentos.	28

3.2	Processo para gerar o conjunto de imagens de BR a partir de uma imagem de AR. (a) Imagem de AR; (b) Imagem apresentando degradações de ruído e blur gaussiano; (c) Imagem reduzida por um $FE= 2$; (d) Versões de BR da imagem com diferentes deslocamentos de sub-píxel.	29
3.3	Representação de um caso ideal para fusão de imagens de BR, sem processo de degradação, em uma imagem de AR, onde a quantidade de imagens é igual ao FE^2 e os deslocamentos das imagens de BR são bem distribuídos no plano de AR. Uma imagem de AR (a) gera um conjunto de 16 imagens de BR (b) com FE igual a 4, sem degradações na imagem e deslocamentos uniformemente distribuídos no plano de AR (c), gerando uma hipótese da imagem de AR idêntica a original, apresentando $REQM = 0$	30
3.4	Representação de um caso para fusão de imagens de BR, sem processo de degradação, em uma imagem de AR, onde a quantidade de imagens é inferior ao FE^2 e os deslocamentos das imagens de BR são bem distribuídos no plano de AR. Uma imagem de AR (a) gera um conjunto de 4 imagens de BR (b) com FE igual a 4, sem degradações na imagem e deslocamentos com distribuição uniforme no plano de AR (c), gerando uma hipótese da imagem de AR apresentando $REQM = 8.79$	31
3.5	Representação de um caso para fusão de imagens de BR, sem processo de degradação, em uma imagem de AR, onde a quantidade de imagens é igual ao FE^2 e os deslocamentos das imagens de BR não são bem distribuídos no plano de AR. Uma imagem de AR (a) gera um conjunto de 16 imagens de BR (b) com FE igual a 4, sem degradações na imagem e deslocamentos aleatoriamente distribuídos no plano de AR (c), gerando uma hipótese da imagem de AR apresentando $REQM = 7.17$	32
3.6	Resultados da avaliação dos fatores de interpolação aplicados ao método de Marcel (1997). a) Média da REQM para o registro de translação; b) Média da REQM para o registro de rotação; Média do TE em segundo; d) Relação entre os erros de registro para a translação e rotação.	34
3.7	Dominância de Pareto no espaço de soluções. Adaptado de Arroyo (2002) . . .	34
3.8	Resultados da avaliação dos fatores de interpolação aplicados ao método de Lucchese (2000). a) Média da REQM para o registro de translação; b) Média da REQM para o registro de rotação; Média do TE em segundo; d) Relação entre os erros de registro para a translação e rotação.	36
3.9	Resultados da avaliação dos valores de raio aplicados ao registro de rotação para o método de Vandewalle (2006). a) Média da REQM para no registro de rotação; b) Média do TE em segundos; c) Relação entre os erros de registro de rotação e TE	37
3.10	Resultados da avaliação dos valores de raio aplicados ao registro de rotação para o método de Vandewalle (2006). a) Média da REQM para no registro de translação; b) Média do TE em segundos; c) Relação entre os erros de registro de translação e TE	38
3.11	Resultados da REQM no registro de rotação dos métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste.	40

3.12	Imagem de número 70 utilizada nos testes. Imagem que originou o conjunto de imagens de BR, que ao ter a rotação estimada pelo método de Keren (1988), no caso de teste 5, apresentou um resultado atípico.	40
3.13	Resultados da REQM no registro de translação no eixo x dos métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste. . .	41
3.14	Resultados da REQM no registro de translação no eixo y dos métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste. . .	42
3.15	Imagens que originaram conjuntos de imagens de BR, que ao terem as translações no eixo x estimadas pelo método de Keren (1988), no caso de teste 5, apresentaram resultado atípico.	42
3.16	Imagens que originaram conjuntos de imagens de BR, que ao terem as translações no eixo y estimadas pelo método de Keren (1988), no caso de teste 5, apresentaram resultado atípico.	43
3.17	Resultados do TE para os métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste.	44
3.18	Resultados dos métodos de reconstrução de uma imagem de teste, para todos os 6 casos de teste (Tabela 3.1), representados em cada linha, sendo a primeira linha o caso de teste 1 e a última o caso de teste 6. Linha 1: FE = 2, qtd. imgs. = 4, res = 240x240; Linha 2: FE = 2, qtd. imgs. = 2, res = 240x240; Linha 3: FE = 4, qtd. imgs. = 16, res = 120x120; Linha 4: FE = 4, qtd. imgs. = 8, res = 120; Linha 5: FE = 6, qtd. imgs. = 36, res = 80x80; Linha 6: FE = 6, qtd. imgs. = 18, res. = 80x80;	46
3.19	Recorte de uma região de interesse de 100x100 píxeis das imagens da Figura 3.18. Todos os 6 casos de teste (Tabela 3.1) estão representados em cada linha, sendo a primeira linha o caso de teste 1 e a última o caso de teste 6. Linha 1: FE = 2, qtd. imgs. = 4, res = 240x240; Linha 2: FE = 2, qtd. imgs. = 2, res = 240x240; Linha 3: FE = 4, qtd. imgs. = 16, res = 120x120; Linha 4: FE = 4, qtd. imgs. = 8, res = 120; Linha 5: FE = 6, qtd. imgs. = 36, res = 80x80; Linha 6: FE = 6, qtd. imgs. = 18, res. = 80x80;	47
3.20	Uma imagem de BR de cada conjunto de teste a ser avaliado. A-C Milanfar (2011); D-F Sunkavalli (2012); G FLIR InfraCamSD.	49
3.21	Resultado da SR para o conjunto A. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.	51
3.22	Resultado da SR para o conjunto B. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.	52
3.23	Resultado da SR para o conjunto C. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.	53

- 3.24 Resultado da SR para o conjunto D. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR. 54
- 3.25 Resultado da SR para o conjunto E. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR. 54
- 3.26 Resultado da SR para o conjunto F. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR. 55
- 3.27 Resultado da SR para o conjunto G. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR. 56

Lista de Tabelas

3.1	Casos de teste propostos para a primeira etapa dos experimentos.	32
3.2	Movimentos de translação (Δx , Δy) e rotação (θ em graus) em relação a imagem de referência para dezesseis imagens de BR no caso de teste 3 e para oito imagens de BR no caso de teste 4.	32
3.3	Resultado da avaliação do fator de interpolação das TF das imagens de BR para o método de Marcel (1997), apresentando: média da REQM para os movimentos horizontais de translação (Δx); média da REQM para os movimentos verticais de translação (Δy); média entre Δx e Δy ; média da REQM para os movimentos de rotação (θ), e; média do TE.	33
3.4	Resultado da avaliação do fator de interpolação das TF das imagens de BR para o método de Lucchese (2000), apresentando: média da REQM para os movimentos horizontais de translação (Δx); média da REQM para os movimentos verticais de translação (Δy); média entre Δx e Δy ; média da REQM para os movimentos de rotação (θ), e; média do TE.	35
3.5	Resultado da avaliação dos valores de raio para estimar a rotação para o método de Vandewalle (2006), apresentando: média da REQM para os movimentos de rotação (θ), e; média do TE.	35
3.6	Resultado da avaliação da quantidade de componentes de baixa frequência para estimar a translação para o método de Vandewalle (2006), apresentando: média da REQM para os movimentos horizontais de translação (Δx); média da REQM para os movimentos verticais de translação (Δy); média entre Δx e Δy , e; média do TE.	37
3.7	Comparação da média da REQM e o seu desvio padrão na estimação dos movimentos de rotação (θ), pelos métodos de registro avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.	39
3.8	Comparação da média da REQM e o seu desvio padrão na estimação dos movimentos de translação (Δx , Δy), pelos métodos de registro avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.	41
3.9	Comparação da média e o seu desvio padrão, do TE gasto para estimar o registro das imagens, pelos métodos de registro avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.	43
3.10	Comparação da média da RSRP e o seu desvio padrão obtido na estimação da imagem de AR pelos métodos de reconstrução avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.	44

3.11	Comparação da média do CCP e o seu desvio padrão obtido na estimação da imagem de AR pelos métodos de reconstrução avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.	45
3.12	Comparação da média do TE, em segundos, e o seu desvio padrão obtido na estimação da imagem de AR pelos métodos de reconstrução avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.	45
3.13	Casos de teste propostos para a primeira etapa dos experimentos.	49

Lista de Abreviaturas e Siglas

AR	-	<i>Alta resolução</i>
BR	-	<i>Baixa Resolução</i>
CCP	-	<i>Coefficiente de Correlação de Pearson</i>
FE	-	<i>Fator de Escala</i>
IBP	-	<i>Iterative Back-Projection</i>
NC	-	<i>Normalized Convolution</i>
REQM	-	<i>Raiz Erro Quadrático Médio</i>
RSR	-	<i>Robust Super-Resolution</i>
RSRP	-	<i>Relação Sinal/Ruído de Pico</i>
SR	-	<i>Super-Resolução</i>
TE	-	<i>Tempo de Execução</i>
TCF	-	<i>Transformada Contínua de Fourier</i>
TDF	-	<i>Transformada Discreta de Fourier</i>
TF	-	<i>Transformadas de Fourier</i>
TV	-	<i>Total Variation</i>
VANT	-	<i>Veículos Aéreos Não-Tripulados</i>

Capítulo 1

Introdução

O objetivo deste capítulo é apresentar quais são os principais aspectos e motivações do estudo desenvolvido. A seguir, são descritos, o problema em questão e a sua relevância, os objetivos do trabalho e a formalização do trabalho.

1.1 Motivação e Relevância

Uma imagem digital pode conter inúmeras informações. Geralmente, as aplicações que utilizam dessas imagens almejam, ou em alguns casos, dependem de imagens maiores resoluções, ou seja imagens de alta resolução (AR). Uma imagem pode ser considerada de AR caso, a cena observada, no domínio contínuo, seja amostrada acima da taxa de Nyquist (Milanfar, 2010). Uma imagem de AR propicia um maior detalhamento do contexto observado, auxiliando o decisor, um usuário ou sistemas de máquina automáticos. Apesar das contínuas pesquisas para desenvolvimento de sensores capazes de capturar imagens de AR, a sua utilização é restrita devido ao elevado custo embutido na fabricação do sensores, inviabilizando algumas aplicações. No entanto, existem métodos de software que permitem combinar as imagens geradas por sistemas de imagens de baixa resolução (BR), mais acessíveis, gerando-se assim uma imagem com resolução aumentada, como por exemplo as técnicas de Super-Resolução (SR) Irani (1991), Zomet (2001), Farsiu (2004) e Maiseli (2014) .

O termo Super-Resolução (SR) pode ser utilizado em diferentes contextos. Ele se relaciona com técnicas que objetivam aprimorar a resolução de sistemas de detecção. No âmbito das imagens digitais, a *resolução espacial* é o foco - assim como no presente trabalho. Resolução espacial se refere a densidade de píxeis em uma imagem, medida em píxeis por unidade de área (Milanfar, 2010). Uma imagem com maior resolução permite que detalhes de uma cena possam ser percebidos com mais clareza. Essa característica possui grande relevância em aplicações de visão computacional. A Figura 1.1 ilustra diferença entre imagens com baixa e alta resolução espacial, entre si.

A nível de hardware, o ganho em resolução espacial pode ser obtido de duas maneiras: diminuindo o tamanho do píxel e aumentando a área do sensor. Uma das restrições para a redução tamanho do píxel é o conjunto óptico. Píxeis cada vez menores são possíveis, porém, as lentes disponíveis não são capazes de representá-los de forma adequada, à medida que seu

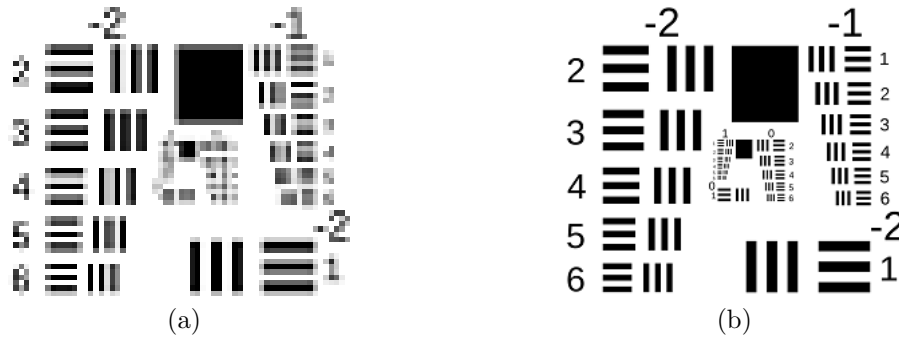


Figura 1.1: Alvo de teste USAF 1951 usado para determinar a resolução espacial de sensores e sistemas de imagens. Imagens com dimensões correspondentes, no entanto, (a) representa uma imagem com baixa resolução espacial, enquanto (b) possui uma resolução quatro vezes maior que (a).

tamanho diminui. De maneira geral, as lentes utilizadas apresentam aberrações geométricas que causam difrações óticas, restringindo a resolução do sensor (Cossairt, 2011). Conseqüentemente, diminuir o tamanho do píxel afeta a sensibilidade à luz (*Daynamic Range*), causando perdas na relação sinal/ruído (*Signal-to-Noise Ratio*) do sensor (Chen, 2000). Além disso, não existe um tamanho ideal para o píxel, a aplicação e os objetivos do projeto são fatores que influenciam diretamente nessa escolha (Hengl, 2006). Por um outro lado, aumentar as dimensões do sensor (aumentar a quantidade de píxeis) acarreta, em um maior consumo de energia aumentando a capacitância do sistema. Um segundo fator é o aumento nos custos de produção e venda, inviabilizando algumas aplicações.

Com as restrições existentes a nível hardware, algumas propostas utilizam das imagens de BR para estimar o seu correspondente em AR. A principal vantagem é que imagens provenientes de sensores comuns de BR, que não possuem especificações complexas e de custos mais baixos, podem ser usados. Alguns métodos de SR utilizam apenas uma imagem de BR (Glasner, 2009; Yang, 2010). Para avaliar a eficiência do estado-da-arte desses métodos, em Yang (2014) são realizados testes de *benchmark*. Uma limitação apresentada é decorrente do fator de escala (FE - taxa de proporção em que a resolução da imagem será aumentada). Para fatores maiores do que 4 os resultados obtidos não são satisfatórios. A capacidade de recuperação dos detalhes nas altas frequências da imagem desses métodos é baixa. A medida em que a riqueza desses detalhes aumenta, o processo para estimar a imagem de AR torna-se mais complexo.

O presente trabalho, no entanto, é centralizado nos métodos de SR que utilizam múltiplas imagens de BR - o termo SR, quando citado, é referente a esse grupo de técnicas. Através de um número maior de imagens mais informação é adicionada ao problema, permitindo uma melhor recuperação dos componentes de alta frequência e diminuindo as degradações causadas pelo processo de captura (Milanfar, 2010). No entanto, essas imagens devem possuir pequenas diferenças de deslocamento não-redundantes e em níveis de sub-píxel. São essas pequenas diferenças que garantem a adição de novas informações relevantes ao problema. A obtenção das imagens pode se dar de várias formas, de uma ou várias câmeras, ou sequências de vídeo. Em um plano comum de alta resolução as imagens são combinadas de acordo com o seu deslocamento

em relação a uma imagem de referência, formando a imagem de AR. A Figura 1.2 ilustra a ideia básica da reconstrução por SR com múltiplas imagens. Se os deslocamentos entre as imagens são conhecidos, ou podem ser estimados com precisão de sub-píxel, é possível a reconstrução de uma imagem por SR (Park, 2003).

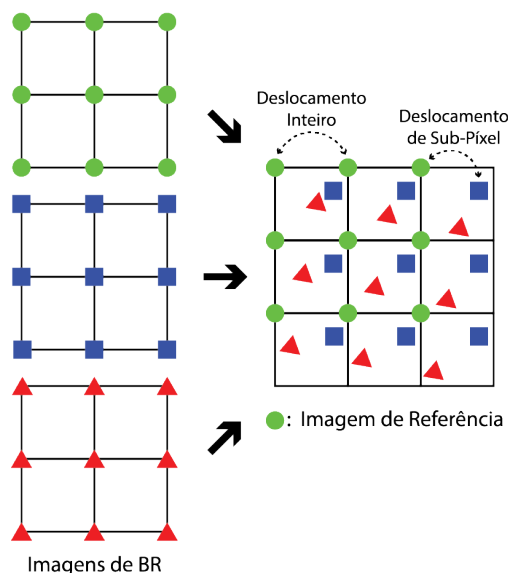


Figura 1.2: Ilustração básica da reconstrução de uma imagem em alta resolução através de múltiplas imagens de BR apresentando deslocamentos a nível de sub-píxel.

Esse método tem sido aplicado em diversos contextos, como objetivo fim, ou, como etapa de pré-processamento para solução de problemas de visão computacional, tais como:

- Imagens Aéreas - imagens capturadas por sistemas de sensoriamento remoto (Zhu, 2012; Pan, 2013) ou Veículos Aéreos Não-Tripulados (VANT) (Heinze, 2010; Wang, 2008)
- Imagens Médicas - exames médicos de imagem como ressonância magnética (Plenge, 2012; Zhang, 2008) e tomografia (Wallach, 2012)
- Imagens Térmicas - em sua maioria os sensores produzem baixa resolução espacial (Panagiotopoulou, 2008; Chikamatsu, 2010)
- Vigilância - Problemas como reconhecimento de faces (Yoshida, 2012; Zou, 2012) , ou melhoria da resolução para melhor análise da cena (Lin, 2005; Zhang, 2010)

Geralmente, o processo SR pode ser dividido em 2 etapas: registro de deslocamento e reconstrução (Vandewalle, 2006; Robinson, 2009). O registro das informações de deslocamento entre as imagens é crucial para o método de SR (Baboulaz, 2009). As frações de deslocamento no conjunto de imagens de BR devem ser inferiores ao tamanho de um píxel, portanto os métodos para estimativa desses deslocamentos devem ser extremamente precisos. O registro é um processo de mapeamento das imagens sobre um mesmo plano de referência. Uma das

imagens é escolhida como referência, e por meio de um modelo de transformação é possível que as imagens sejam alinhadas geometricamente.

A fase de reconstrução pode ser dividida em interpolação e restauração. A interpolação é o processo de alinhar as imagens de BR, com precisão de sub-píxel, e combiná-las em um plano de AR (Milanfar, 2010). Uma vez que os deslocamentos entre as imagens de BR são arbitrários, nem sempre a imagem de AR irá se mostrar uniformemente espaçada no plano de AR. Portanto, processo de interpolação de imagens de BR, não uniformemente espaçadas, é necessária para obtenção de uma imagem de AR uniformemente espaçada (Park, 2003). A imagem resultante desse processo pode apresentar imperfeições, devido aos equívocos no registro das imagens ou a falta de imagens relevantes para o problema. Por isso, a restauração tem como objetivo tratar as incertezas do processo oferecendo uma imagem de AR com mais qualidade.

Este trabalho busca estudar a SR com múltiplas imagens, avaliando o desempenho de diferentes métodos de registro de imagens e reconstrução, aplicados a diferentes cenários, analisando os resultados para uma determinar quais métodos melhor se aplicam ao um contexto em questão.

1.2 Definição do Problema de Super-Resolução

As abordagens utilizadas pelos métodos de SR procuram estabelecer um processo inverso ao processo de captura de um sistema de imagem. As câmeras digitais usam de um modelo que captura observações do mundo real gerando imagens digitais. No processo de captura perdas naturais ocorrem no sinal observado.

A Figura 1.3 exemplifica o processo de captura de uma imagem. A primeira etapa no processo de captura é a discretização do sinal contínuo (a). Para que o sistema de captura consiga recuperar uma imagem contínua, a frequência de amostragem deve atender o teorema de Nyquist. O teorema afirma que a taxa de amostragem deve ser maior que o dobro da frequência de corte do conjunto óptico. No entanto, devido as limitações dos sistemas de detecção, as imagens frequentemente apresentam certo “serrilhamento”(aliasing). Para as técnicas de SR a imagem de AR a ser recuperada respeita os critérios de Nyquist. A informação de entrada da câmera são múltiplas imagens (ou quadros no caso de um vídeo) com possíveis deslocamentos entre si (b). Devido as limitações ópticas, quando as observações capturadas passam pelas lentes da câmera estas sofrem efeitos de “borrão”(blur) (c). Ao atingir o sensor as imagens tem a sua resolução reduzida, além dos efeitos causados pelo ruído. Finalmente, depois de todo processo, as imagens são apresentadas à saída do sistema.

Considerando uma imagem de AR $\mathbf{f} = [f_1, f_2, \dots, f_N]^T$ amostrada acima da taxa de Nyquist, de tamanho $L_1 N_1 \times L_2 N_2$ representada por um vetor unidimensional, rearranjado em ordem lexicográfica, onde, f é um píxel, N a quantidade total de píxeis na imagem com $N = L_1 N_1 \times L_2 N_2$ e, L_1 e L_2 sendo, respectivamente, os fatores de redução de resolução horizontais e verticais, as amostras de BR geradas possuem tamanho $N_1 \times N_2$. A k -ésima observação pode ser descrita por, $\mathbf{g}_k = [g_{k,1}, g_{k,2}, \dots, g_{k,M}]^T$, onde $k = 1, 2, \dots, p$, p é a quantidade de imagens de BR, com M píxeis sendo, $M = N_1 \times N_2$. Tomando \mathbf{f} constante durante a aquisição das imagens de BR, exceto pela degradação e pelos deslocamentos, o modelo de observação pode ser representado por (Park, 2003)

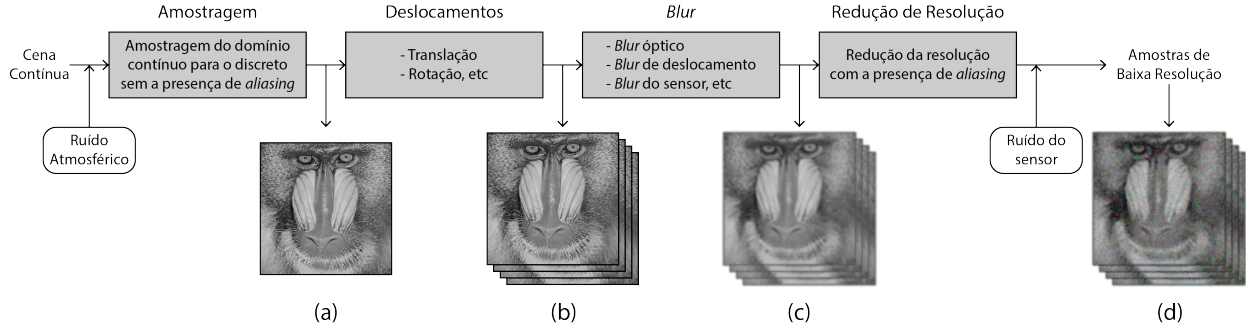


Figura 1.3: Modelo de observação de um sistema de captura de imagens relacionando uma cena em alta resolução e as imagens de BR geradas pela câmera (adaptado de Milanfar (2010)).

$$\mathbf{g}_k = D_k B_k M_k \mathbf{f} + n_k, \quad \text{para } 1 \leq k \leq p \quad (1.1)$$

onde, M_k , de tamanho $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, é a matriz de encapsulamento, B_k , de tamanho $L_1 N_1 L_2 N_2 \times L_1 N_1 L_2 N_2$, a matriz dos “borrões” (*blur*), D_k , de tamanho $(N_1 N_2)^2 \times L_1 N_1 L_2 N_2$, matriz de redução de resolução e, n_k o vetor que reproduz o ruído do processo.

Geralmente, os deslocamentos contidos no processo de captura podem ser variados - translação, rotação ou transformações afins. Esses deslocamentos são codificados pela matriz M_k . Os “borrões” causados pelo conjunto óptico são modelados em B_k e os operadores de redução de resolução, que definirão as dimensões da observação, são definidos por D_k .

O grande desafio da SR se deve ao fato de que a matrizes representadas por D_k , B_k e M_k são muito esparsas, além desse sistema linear ser tipicamente um problema mal condicionado. Em sistemas reais de captura essas matrizes são desconhecidas (Milanfar, 2010). Portanto, o método de SR deve ser capaz de estimar, de maneira eficaz, cada uma dessas matrizes, e com base nas observações de BR disponíveis, buscar reverter da melhor forma o processo de captura, reproduzindo a cena de AR observada.

1.3 Abordagem

As imagens de BR são obtidas através amostras de uma cena em um domínio contínuo de AR. No entanto, apenas com essas amostras de BR, não é possível estabelecer, com exatidão, a versão discreta de AR da cena. Não se sabe qual foi a cena que proveu informações para a formação das imagens de BR. Este é o principal desafio da SR. Aproximar-se do resultado ótimo, do *ground truth* do problema.

Com o intuito de avaliar as técnicas de registro e reconstrução, se faz necessário experimentos com imagens de BR artificialmente simuladas. As imagens de BR I_{Bk} , com k sendo o número de imagens de BR, são provenientes de uma imagem I_A , que é definida como a solução ótima do problema (o *ground truth*). Atribuindo filtros de *blur* e ruído nas imagens, além de diferentes operações de transformação em I_A é possível simular um sistema de captura de imagens e gerar um conjunto de imagens de BR, I_{Bk} . Assim, a SR pode ser aplicada em I_{Bk} com o objetivo de recuperar I_A .

A aplicação da SR é possível através da combinação de uma técnica de registro Reg com uma técnica de reconstrução Rec . Ao combinar essas técnicas e aplicá-las ao conjunto I_{B_k} é possível gerar uma imagem I_{Ah} , ou seja, uma hipótese de I_A . Ao comparar I_{Ah} , o resultado proveniente da SR, com I_A é possível avaliar a acurácia da SR através da combinação $[Reg, Rec]$. Assim, pretende-se, no trabalho, definir uma combinação $[Reg, Rec]$ que permita a obtenção de imagens de AR de qualidade.

Uma vez que os experimentos são provenientes de simulações, é importante que a combinação definida seja aplicada para problemas com imagens não simuladas, ou seja, conjuntos de imagens de BR provenientes de sistemas de captura. Assim, é possível validar, com dados reais, a escolha da combinação definida nos experimentos simulados.

1.4 Objetivos

1.4.1 Objetivos Gerais

O principal objetivo do presente trabalho é estudar a técnica de Super-Resolução a partir de múltiplas imagens de baixa resolução, bem como os métodos aplicados nas etapas de registro e reconstrução, aplicando-os a diferentes cenários e avaliando a sua acurácia na construção de imagens de alta resolução.

1.4.2 Objetivos Específicos

Os objetivos específicos desse trabalho são:

1. Estudar a Super-Resolução com múltiplas imagens de baixa resolução.
2. Investigar métodos de registro de imagens com precisão de sub-píxel.
3. Investigar métodos para reconstrução de imagens de alta resolução por meio de imagens com baixa resolução.
4. Avaliar os métodos de registro e reconstrução investigados em cenários favoráveis a utilização da SR.
5. Definir a melhor combinação de uma técnica de registro e uma técnica reconstrução, dentre os métodos investigados, para aplicação da SR.
6. Validar a combinação escolhida por meio de aplicação em problemas de SR.

1.5 Contribuição

A SR com múltiplas imagens é uma solução para o aumento de resolução de imagens. Diversas são as técnicas que englobam as duas fases do processo de SR, o registro (Deshmukh, 2011) e a reconstrução (Nasrollahi, 2014). A investigação de uma combinação que reúne uma

técnica de registro e uma de reconstrução, e produz uma imagem de AR com a acurácia desejada é um processo custoso.

O presente trabalho, vem então, trazer uma comparação de 4 técnicas de registro e 5 técnicas de reconstrução, para auxiliar a escolha das respectivas técnicas possibilitando a sua combinação e aplicação à SR de uma maneira mais ágil, trazendo maiores informações a respeito das técnicas avaliadas.

Espera-se uma definição contundente da melhor combinação de técnicas para prover um caminho melhor direcionado para a utilização da SR.

1.6 Organização do Texto

A parte restante deste trabalho está organizada da seguinte forma: o Capítulo 2 apresenta a revisão de literatura do trabalho. Os métodos de registro de imagens de BR, assim como, os métodos de reconstrução, abordados pelo trabalho, são apresentados descrevendo as suas características e metodologia usada. No Capítulo 3 os experimentos realizados são demonstrados. Primeiramente, são realizados experimentos baseados em imagens de BR geradas artificialmente a partir de uma imagem de AR. Os métodos de registro são avaliados apontando aquele que obteve melhores resultados. Logo em seguida, os métodos de reconstrução são experimentados definindo, qual dentre os métodos investigados possui a melhor estimativa da imagem de AR a ser construída. Selecionando os métodos com melhores resultados nas etapas de registro e reconstrução, esses métodos combinados são aplicados a conjuntos de imagens de BR com deslocamentos e processo de amostragem desconhecidos, com o objetivo de validar a combinação escolhida. Por fim, no Capítulo 4 é apresentada as conclusões obtidas na realização do trabalho.

Capítulo 2

Revisão de Literatura

A proposta do presente capítulo é revisar a bibliografia que engloba o trabalho. Primeiramente, as técnicas de registro de imagens abordadas são descritas. Logo após, os cinco métodos para reconstruir uma imagem de AR a partir de múltiplas imagens de BR, que se utilizam das informações provenientes do registro, são expostos.

2.1 Introdução

O objetivo da SR é construir uma imagem de AR através de um conjunto de imagens de BR. Para isso, é necessário que as imagens de BR tenham uma característica essencial, diferenças de deslocamentos a níveis de sub-píxel entre si (Park, 2003). Conhecendo o deslocamento entre imagens é possível adicionar informação relevante ao problema de aumento de resolução, e consequentemente, melhorar o resultado da imagem de AR, pela aplicação da SR. A Figura 2.1 mostra como os deslocamentos de sub-píxel influenciam a formação da imagem de AR. A partir de um conjunto contendo imagens de BR (a) com deslocamentos em níveis de sub-píxel entre si, é possível, conhecendo os valores precisos dos deslocamentos (b), a sua interpolação (c) gerando uma imagem com resolução aumentada (d).

Caso as imagens de BR possuam deslocamentos inteiros entre si, estes não contribuirão com informações relevantes ao problema. Imagens de BR com deslocamentos inteiros entre si estarão parcialmente sobrepostas na imagem de AR, o mesmo que informação redundante (Park, 2003). Esse fato é exemplificado na Figura 2.2. Diferentemente da Figura 2.1, uma das imagens de BR possui deslocamento inteiro em relação a imagem de referência (Figura 2.2b). Como os deslocamentos são conhecidos, os píxeis sobrepostos indicam que estes possuem a mesma informação. Assim, a imagem de BR com deslocamento inteiro, não contribuirá para formar a imagem de AR (Figura 2.2d). Comparando as Figuras 2.1d e 2.2d é possível perceber que a quantidade de informações obtidas na Figura 2.2d é inferior a Figura 2.1d. A tendência é que, com mais informações relevantes o processo de SR produza uma imagem de AR melhor.

No processo de formação das imagens, os tipos de deslocamentos podem ser causados por transformações de, translação, rotação, escala, afins, etc. A precisão no registro é difícil de ser alcançada quando é preciso estimar um grande número de parâmetros de transformação, o que pode potencializar os erros de registro (Hardie, 2007). Devido a precisão de sub-píxel necessária

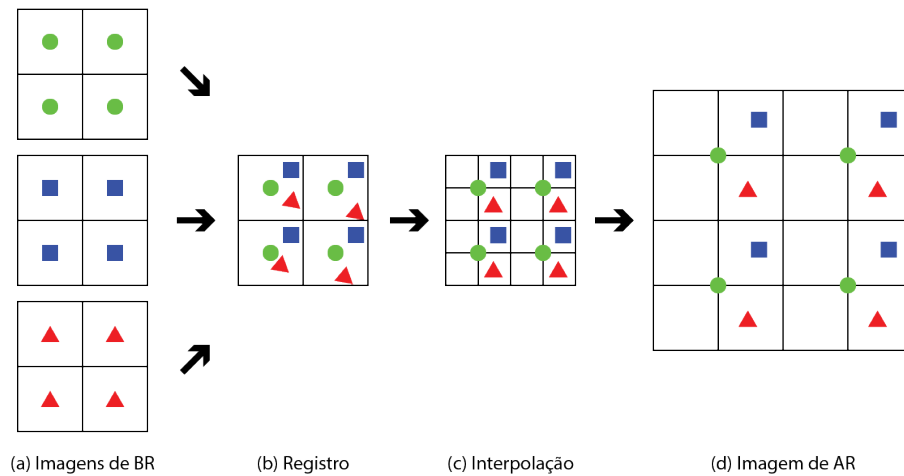


Figura 2.1: Interpolação de múltiplas imagens de BR, com deslocamentos de sub-píxel, em um imagem com resolução aumentada. (a) conjunto de imagens de BR; (b) as imagens organizadas de acordo com os seus deslocamentos; (c) combinação das imagens em um plano de resolução aumentada; (d) resultado do processo de interpolação.

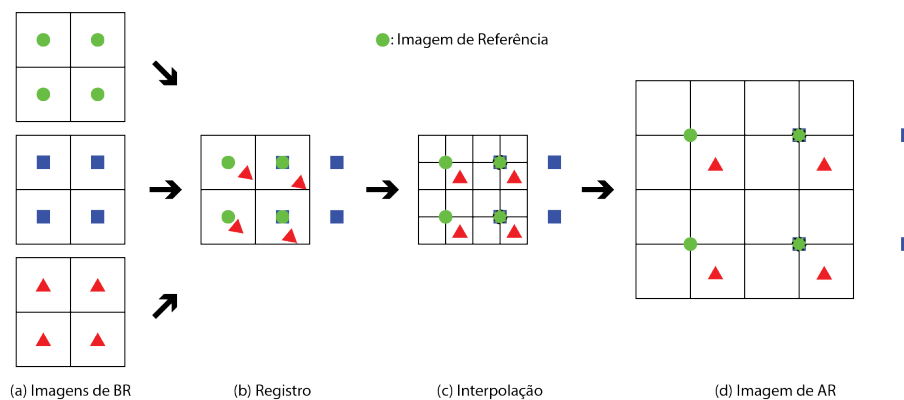


Figura 2.2: Interpolação de múltiplas imagens de BR, com deslocamentos inteiros e de sub-píxel, em um imagem com resolução aumentada. (a) conjunto de imagens de BR; (b) as imagens organizadas de acordo com os seus deslocamentos; (c) combinação das imagens em um plano de resolução aumentada; (d) resultado do processo de interpolação.

no registro das imagens de BR, o presente trabalho considera apenas as transformações de translação e rotação, assim como em, Tian (2010), Gao (2011) e Karch (2015).

Além do registro preciso dos deslocamentos entre as imagens de BR, o número de imagens do conjunto de entrada e o FE escolhido irão influenciar diretamente na performance da SR (Baker, 2002). Para que seja possível bons resultados, a quantidade de imagens disponíveis deve estar relacionada ao FE, ou seja, quanto maior o FE escolhido, mais imagens se fazem necessárias. Observando a Figura 2.3 é possível perceber que, considerando um ambiente ideal onde as imagens de BR tem deslocamentos bem distribuídos, a quantidade de imagens BR necessária para preencher a imagem de AR é igual ao FE². Quando essa relação é estabelecida, a quantidade de pixels do conjunto de BR é igual a quantidade de pixels da imagem de AR

geradora. Em experimentos realizados por Lin (2004), chegou-se a conclusão que, quando o FE escolhido é de ordem inteira, a quantidade suficiente de imagens é igual FE^2 . Porém, ao selecionar um FE de ordem real, a quantidade ideal passa a ser $4FE^2$.

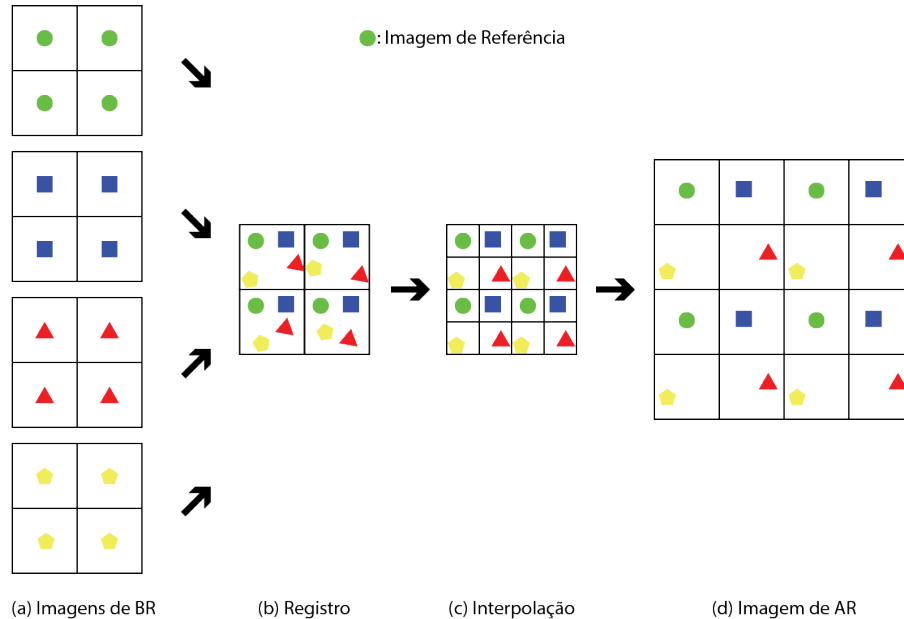


Figura 2.3: Interpolação de múltiplas imagens de BR, com deslocamentos de sub-píxel, em um imagem com resolução aumentada, com a quantidade de píxeis do conjunto de BR igual a quantidade de píxeis da imagem de AR. (a) conjunto de imagens de BR; (b) as imagens organizadas de acordo com os seus deslocamentos; (c) combinação das imagens em um plano de resolução aumentada; (d) resultado do processo de interpolação.

Uma vez finalizado o registro das imagens os valores estimados são fornecidos como entrada ao método de reconstrução. A imagem de AR será reconstruída a partir das imagens de BR disponíveis. Os primeiros estudos relacionados à utilização de múltiplas imagens de BR foram introduzidos por Tsai (1984), com uma abordagem baseada no domínio da frequência com três princípios:

- i - A propriedade de deslocamento da transformada de Fourier.
- ii - A relação de “serrilhamento” (*aliasing*) entre a Transformada Contínua de Fourier (TCF) da imagem original de AR e as Transformadas Discretas de Fourier (TDF) das imagens de BR observadas.
- iii - A suposição de que a imagem original de AR tem um limite de banda.

Essas propriedades possibilitam a formulação de um sistema de equações relacionando os coeficientes da TDF das observações de BR para gerar uma amostra da TCF de uma imagem desconhecida.

Várias extensões foram propostas a partir desta abordagem, considerando alterações no modelo de observação, efeitos de *blur*, ruído, e adicionando termos de regularização (Nasrollahi,

2014). Porém, apesar da sua eficiência computacional, os métodos baseados no domínio da frequência possuem limitações em relação aos modelos de degradação de imagem mais complexos. Portanto, trabalhos posteriores tiveram a sua atenção voltada para abordagens no domínio espacial problema (Milanfar, 2010).

Os métodos espaciais de reconstrução baseiam-se dos efeitos de *aliasing* presentes nas imagens de BR com o objetivo de aumentar a resolução da imagem, por meio de um modelo de observação, que busca simular o processo ao qual as imagens de BR foram formadas (Nasrollahi, 2014). Um dos primeiros métodos desenvolvidos com essa abordagem espacial foi o proposto por Irani (1991), o *Iterative Back Projection* (IBP), ou retroprojeção iterativa (descrito com mais detalhes na Sessão 2.3.1). No IBP a imagem de AR é estimada a partir da retroprojeção do erro entre, as imagens de BR simuladas provenientes da estimativa de AR e as imagens de BR amostradas. Esse processo é repetido de forma iterativa, buscando a minimização do valor de erro.

O IBP pode ser classificado como uma abordagem por máxima verossimilhança (*Maximum-Likelihood* - ML). O objetivo da ML é maximizar a probabilidade das imagens de BR serem resultados de uma cena em comum (Pham, 2006b). Porém, os métodos baseados na ML apresentam problemas mal condicionado, ou seja, sensíveis a ruído e erros na fase de registro, além da possibilidade de não existir uma solução única para o problema (Nasrollahi, 2014). Para solucionar esses problemas é necessário que uma informação adicional seja inserida no problema, com o objetivo de restringi-lo. Para isso, surgiu a classe dos métodos com uma abordagem pelo máximo *a posteriori* (*maximum a posteriori* - MAP). Nesta classe de métodos, um termo age como operador de regularização, trazendo estabilidade ao problema (Park, 2003). Na literatura diferentes tipos de operadores de regularização tem sido utilizados. No presente trabalho algumas técnicas que utilizam desses operadores de regularização são avaliadas a seguir (Sessão 2.3.2, Sessão 2.3.4 e Sessão 2.3.5).

Nas Sessões abaixo são apresentados cada um dos métodos investigados no trabalho. Ao todo são 4 métodos de registro e 5 métodos de reconstrução. A escolha dos métodos foi realizada pela viabilidade, favorecendo a sua utilização na aplicação da SR. Os métodos investigados não representam o estado-da-arte de cada uma das áreas. Porém, como esses são de fácil acesso e aplicação ampliam a possibilidade do seu uso.

2.2 Métodos de Registro

2.2.1 Keren

O método desenvolvido por Keren (1988) é uma abordagem espacial para o registro de imagens. O método é capaz de estimar as transformações de translação e rotação entre imagens de BR. O método realiza a estimativa com base nas expansões da série de Taylor. Com o intuito de aprimorar a precisão e o custo computacional do método, um esquema piramidal de escalas é utilizado de forma iterativa. O registro é primeiramente realizado com as imagens em escalas reduzidas. Iterativamente, o método propaga os resultados para escalas superiores, até que a resolução original da imagem seja atingida.

Considerando uma imagem \mathbf{g}_0 e \mathbf{g}_1 sendo uma versão de \mathbf{g}_0 apresentando diferenças de

deslocamento horizontal (a), vertical (b) e um ângulo de rotação (θ) com origem no eixo, a relação entre elas pode ser expressa por

$$\mathbf{g}_1(x, y) = \mathbf{g}_0(x \cos(\theta) - y \sin(\theta) + a, y \cos(\theta) + x \sin(\theta) + b). \quad (2.1)$$

Expandindo $\sin(\theta)$ e $\cos(\theta)$ aos primeiros dois termos da série de Taylor tem-se

$$\mathbf{g}_1(x, y) \approx \mathbf{g}_0 \left(x + a - y\theta - \frac{x\theta^2}{2}, y + b + x\theta - \frac{y\theta^2}{2} \right). \quad (2.2)$$

Utilizando novamente a série de Taylor para expandir g_0 ao primeiro termo de sua própria série, a seguinte equação de primeira ordem pode ser observada

$$\mathbf{g}_1(x, y) \approx \mathbf{g}_0(x, y) + \left(a - y\theta - \frac{x\theta^2}{2} \right) \frac{\partial \mathbf{g}_0}{\partial x} + \left(b - x\theta - \frac{y\theta^2}{2} \right) \frac{\partial \mathbf{g}_0}{\partial y}. \quad (2.3)$$

A função de erro entre \mathbf{g}_0 e \mathbf{g}_1 após a rotação por θ e a translação por a e b pode ser aproximada por

$$E(a, b, \theta) = \sum \left[\mathbf{g}_0(x, y) + \left(a - y\theta - \frac{x\theta^2}{2} \right) \frac{\partial \mathbf{g}_0}{\partial x} + \left(b - x\theta - \frac{y\theta^2}{2} \right) \frac{\partial \mathbf{g}_0}{\partial y} - \mathbf{g}_1(x, y) \right]^2. \quad (2.4)$$

onde a somatória se dá sobre a área de sobreposição entre \mathbf{g}_0 e \mathbf{g}_1 .

Ao observar o mínimo de $E(a, b, \theta)$ através do cálculo das derivadas parciais de a , b e θ , igualando-as a zero e rejeitando os termos não lineares, é possível obter o seguinte sistema de equações lineares

$$\begin{aligned} \left[\sum \left(\frac{\partial \mathbf{g}_0}{\partial x} \right)^2 \right] a + \left[\sum \frac{\partial \mathbf{g}_0}{\partial x} \frac{\partial \mathbf{g}_0}{\partial y} \right] b + \left[\sum R \frac{\partial \mathbf{g}_0}{\partial x} \right] \theta &= \sum \frac{\partial \mathbf{g}_0}{\partial x} (\mathbf{g}_0 - \mathbf{g}_1) \\ \left[\sum \frac{\partial \mathbf{g}_0}{\partial x} \frac{\partial \mathbf{g}_0}{\partial y} \right] a + \left[\sum \left(\frac{\partial \mathbf{g}_0}{\partial y} \right)^2 \right] b + \left[\sum R \frac{\partial \mathbf{g}_0}{\partial y} \right] \theta &= \sum \frac{\partial \mathbf{g}_0}{\partial y} (\mathbf{g}_0 - \mathbf{g}_1) \end{aligned} \quad (2.5)$$

$$\left[\sum R \frac{\partial \mathbf{g}_0}{\partial x} \right] a + \left[\sum R \frac{\partial \mathbf{g}_0}{\partial y} \right] b + \left[\sum R^2 \right] \theta = \sum R (\mathbf{g}_0 - \mathbf{g}_1)$$

onde R é uma abreviatura de $x \frac{\partial \mathbf{g}_0}{\partial y} - y \frac{\partial \mathbf{g}_0}{\partial x}$ e o somatório acontece na área de sobreposição das imagens.

Tomando como base as aproximações feitas em (2.3), a expressão é verdadeira apenas quando os valores de (a, b, θ) são pequenos. Portanto, para obter uma acurácia a nível de sub-píxel um processo iterativo é realizado. Ao solucionar as equações, a imagem \mathbf{g}_1 é repetidamente deslocada por meio de (2.1). A cada iteração ocorre o acúmulo dos valores de (a, b, θ) (deslocando \mathbf{g}_1) até que um determinado número de iterações seja atingido, ou a solução resultante seja bem pequena. Uma vez que \mathbf{g}_0 é sempre a mesma, os parâmetros das funções são calculados uma única vez, alterando a cada iteração apenas os valores escalares.

Com o intuito de aumentar a robustez e a velocidade do método, um esquema de pirâmides gaussianas é utilizado. Neste esquema, a imagem original de tamanho $N \times N$ passa por

filtros gaussianos que vão reduzindo as dimensões pela metade, $N/2 \times N/2$. Primeiramente, os parâmetros de movimento são calculados para uma imagem pequena (geralmente 64×64). Os valores encontrados são interpolados para a escala superior seguinte, e corrigidos por uma ou duas iterações. O processo de interpolação e correção ocorre até que as dimensões originais da imagem sejam atingidas. A complexidade computacional obtida por meio do processo piramidal é a mesma obtida com duas iterações utilizando as imagens originais.

2.2.2 Marcel

O método para registro de imagens desenvolvido por Marcel (1997) se baseia no domínio da frequência através de Transformada de Fourier (TF). O método foi primeiramente desenvolvido para estimar os deslocamentos translacionais, observando a fase das TF das imagens. A extensão do método para determinar a rotação parte do princípio de que, uma vez que o plano de frequência é transformado em coordenadas polares, a rotação entre as imagens se comporta como um deslocamento de translação. Portanto, o mesmo método pode registrar a translação e rotação entre as imagens.

Considerando $\mathbf{g}_0(\mathbf{x})$, para $\mathbf{x} = (x, y) \in \mathbb{R}^2$, uma imagem no domínio espacial, a sua versão com deslocamento de translação $\mathbf{g}_1(\mathbf{x})$ é obtida por

$$\mathbf{g}_1(\mathbf{x}) = \mathbf{g}_0(\mathbf{x} - t). \quad (2.6)$$

onde t é o operador de translação correspondente. A transformação de $\mathbf{g}_0(\mathbf{x})$ do domínio espacial para o domínio da frequência ocorre pela aplicação de TF

$$\begin{aligned} G_0(\mathbf{u}) &= \mathcal{F}[\mathbf{g}_0(\mathbf{x})] \\ &= \int_{\mathbb{R}^2} \mathbf{g}_0(\mathbf{x}) e^{-2j\pi\mathbf{u}^T \mathbf{x}} d\mathbf{x}. \end{aligned} \quad (2.7)$$

Portando a representação da imagem deslocada $\mathbf{g}_1(\mathbf{x})$ no domínio da frequência é obtida por

$$\begin{aligned} G_1(\mathbf{u}) &= \mathcal{F}[\mathbf{g}_0(\mathbf{x} - t)] \\ &= G_0(\mathbf{u}) e^{-2j\pi\mathbf{u}^T t}. \end{aligned} \quad (2.8)$$

Sendo $\Psi(\mathbf{u})$ a razão entre os espectros (espectro de potência cruzado ou espectro cruzado) de \mathbf{g}_0 e \mathbf{g}_1 , Ψ pode ser calculada por

$$\begin{aligned} \Psi(\mathbf{u}) &= \frac{G_1(\mathbf{u})}{G_0(\mathbf{u})} \\ &= e^{-2j\pi\mathbf{u}^T t}. \end{aligned} \quad (2.9)$$

O cálculo do espectro de potência cruzado Ψ gera uma função exponencial complexa que é compreendida por um vetor correspondente ao vetor de translação $-t$. Os valores das variáveis de translação podem ser estimados através do ponto de máximo Ψ .

Utilizando essa observação, Marcel (1997) estenderam a abordagem para estimar a rotação entre as imagens. Para isso, o plano de coordenadas de frequência é convertido para coordenadas polares e a magnitude dos espectros é analisada.

Acrescentando um ângulo de rotação θ a \mathbf{g}_1 , a relação entre \mathbf{g}_0 e \mathbf{g}_1 com deslocamento de translação e movimento de rotação pode ser definida por

$$\begin{aligned}\mathbf{g}_1(x) &= \mathbf{g}_0(R(\theta)^{-1}\mathbf{x} - t) \\ &= \mathbf{g}_0(R(\theta)^{-1}(\mathbf{x} - R(\theta)t)).\end{aligned}\tag{2.10}$$

A definição dessa relação entre as imagens no domínio da frequência é obtida por

$$G_1(\mathbf{u}) = G_0(R(\theta)^{-1}\mathbf{u})e^{-2j\pi\mathbf{u}^T R(\theta)t}\tag{2.11}$$

onde

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.\tag{2.12}$$

Com as magnitudes das TF das imagens é possível estimar a rotação independente da translação. Os deslocamentos espaciais de translação afetam apenas os valores de fase das TF. Portanto, a relação das magnitudes das imagens no domínio da frequência pode ser estabelecida como

$$|G_1(\mathbf{u})| = |G_0(R(\theta)^{-1}\mathbf{u})|.\tag{2.13}$$

É possível afirmar que $|G_1(\mathbf{u})|$ é uma versão rotacionada de $|G_0(\mathbf{u})|$ sobre o mesmo ângulo θ do domínio espacial, pois $|G_1(\mathbf{u})|$ e $|G_0(\mathbf{u})|$ não são dependentes dos deslocamentos de translação. Assim, é possível estimar a rotação separadamente da translação.

Transformando o plano de frequência em coordenadas polares a transformação de rotação é considerada como uma translação (circular) sobre o ângulo de rotação. Considerando, \mathbf{p} um ponto no plano de coordenadas polares no domínio da frequência, onde $|G_{0_{pol}}(\mathbf{p})|$ e $|G_{1_{pol}}(\mathbf{p})|$ representam, respectivamente, $|G_1(\mathbf{u})|$ e $|G_0(\mathbf{u})|$ em coordenadas polares, calcula-se o espectro de potência cruzado ψ_{pol} por

$$\begin{aligned}\psi_{pol}(\mathbf{p}) &= \frac{|G_{1_{pol}}(\mathbf{p})|}{|G_{0_{pol}}(\mathbf{p})|} \\ &= e^{-2j\pi\mathbf{u}^T \mathbf{p}t_{pol}}\end{aligned}\tag{2.14}$$

onde o vetor de onda do espectro ψ_{pol} contém o vetor de translação $-t_{pol}$ que representa o ângulo de rotação entre as imagens.

2.2.3 Lucchese

O método proposto por Lucchese (2000), assim como Marcel (1997) também se baseia no domínio da frequência utilizando-se das propriedades da magnitude dos espectros de frequência e de coordenadas polares. Estudando a relação de translação entre sinais, Lucchese (2000) verificaram que, a diferença entre um sinal e o espelhamento da sua versão transladada produz um único ponto de passagem pelo zero, pelo qual a translação entre estes pode ser obtida. Com isso, estendendo essa ideia de diferença para as magnitudes das TF (2.13) é possível estimar a rotação entre as imagens. A diferença $\Delta(\mathbf{u})$ entre $|G_0(\mathbf{u})|$ a reflexão de $|G_1(\mathbf{u})|$ é definida por

$$\begin{aligned}
\Delta(\mathbf{u}) &= \frac{|G_0(\mathbf{u})|^2}{G_0^2(0)} - \frac{|G_1(H\mathbf{u})|^2}{G_1^2(0)} \\
&= \frac{|G_0(\mathbf{u})|^2}{G_0^2(0)} - \frac{|G_0(R(\theta)^{-1}H\mathbf{u})|^2}{G_0^2(0)}
\end{aligned} \tag{2.15}$$

onde $H = \begin{bmatrix} \pm 1 & 0 \\ 0 & \mp 1 \end{bmatrix}$ executa a reflexão relativo ao eixo \mathbf{u}_x ou \mathbf{u}_y com base nos sinais escolhidos na diagonal principal. O locus, onde $\Delta(\mathbf{u}) = 0$, produz um par de linhas ortogonais de passagem pela origem, onde o ângulo formado por essas linhas com os eixos é igual a metade do ângulo de rotação entre as imagens.

A rotação é estabelecida em três fases. O primeiro estágio tem o objetivo de encontrar o ponto de acumulação da coordenada angular que define o declive das linhas. Isso é possível através da transformação das coordenadas cartesianas para coordenadas polares, que permite construir um histograma $\theta = \mathcal{H}(-2\varphi)$ onde o ponto de máximo representa o ângulo de rotação. As duas etapas seguintes são etapas de refinamento da solução. Um novo histograma é construído em uma resolução maior, limitando-se à vizinhança da solução encontrada. A última etapa usa regressão linear baseada em mínimos quadrados para minimização do erro quadrático médio de φ .

Após o cálculo da rotação, a imagem rotacionada é compensada pelo ângulo de rotação encontrado e um método de correlação de fase (como o abordado por Marcel (1997)) é utilizado para o cálculo de translação entre as imagens.

2.2.4 Vandewalle

Assim como nos métodos propostos por Marcel (1997) e Lucchese (2000), o trabalho de Vandewalle (2006) utiliza o domínio da frequência com base em coordenadas polares, estimando as movimentações decorrentes de transformações planares de forma separada, primeiramente a rotação e logo após a translação.

Considerando $G_0(\mathbf{u})$ uma imagem de referência no domínio da frequência e $G_1(\mathbf{u})$ a sua versão rotacionada, a estimativa de rotação se dá através da análise das magnitudes de $G_0(\mathbf{u})$ e $G_1(\mathbf{u})$, representadas por, $|G_0(\mathbf{u})|$ e $|G_1(\mathbf{u})|$ (Figura 2.4 (a) e (b)). O ângulo de rotação entre as imagens, corresponde ao ângulo θ em que, $|G_0(\mathbf{u})|$ e $|G_1(R_\theta\mathbf{u})|$ apresentam a correlação máxima. Através das coordenadas polares, o conteúdo de frequência h é calculado em função do ângulo α pela integração através de linhas radiais:

$$h(\alpha) = \int_{\alpha+\Delta\alpha/2}^{\alpha-\Delta\alpha/2} \int_0^\infty |G(r, \theta)| dr d\theta. \tag{2.16}$$

Na prática $|G(r, \theta)|$ é um sinal discreto. Portanto, $h(\alpha)$ é calculado como a média dos valores dentro de um retângulo delimitado por $\alpha - \Delta\alpha/2 < \theta < \alpha + \Delta\alpha/2$. A precisão definida é de 0.1, desta forma, $h(\alpha)$ é calculado variando α em valores de 0.1 graus. Para que um número similar de sinais seja possível para cada valor de α , a média é avaliada dentro de um círculo com raio $r < \rho$ (ρ é o raio da imagem). Finalmente, com o objetivo de evitar as interferências dos valores de baixa frequência, esses são descartados sujeitando-se a um limiar, onde $r < \rho\epsilon$, com $\epsilon = 0.1$. A função $h(\alpha)$ aplicada à $|G_0(\mathbf{u})|$ e $|G_1(\mathbf{u})|$ (Figura 2.16 (c)) resulta em uma

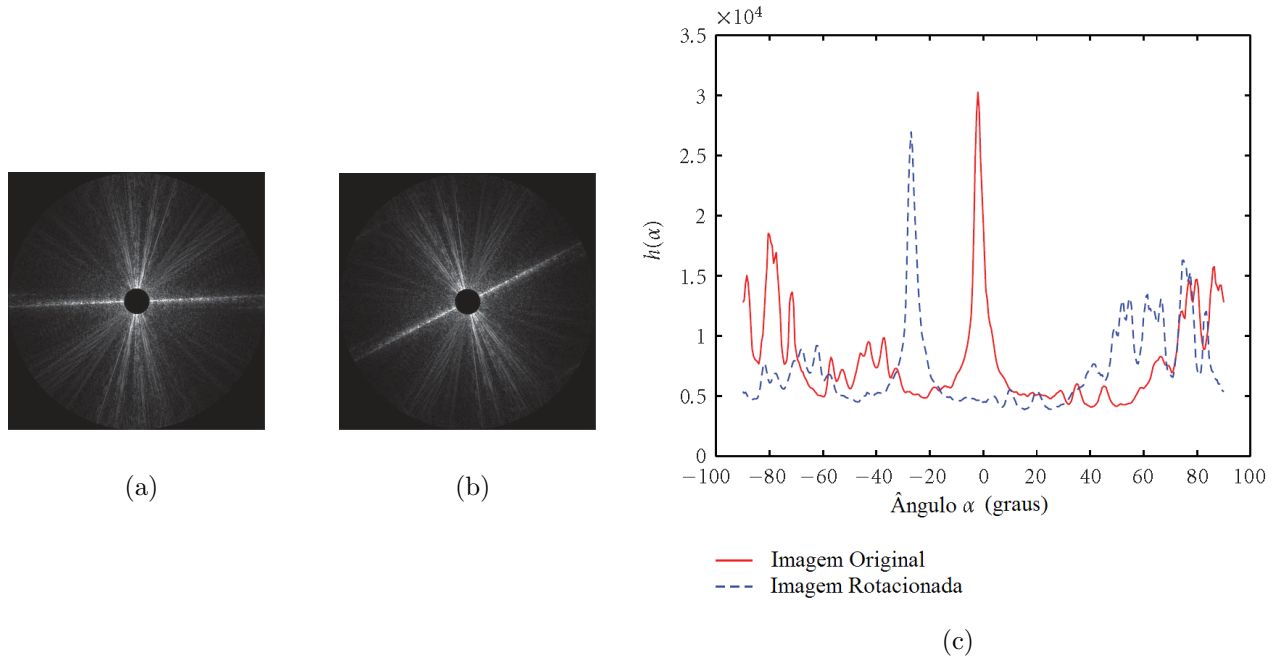


Figura 2.4: Estimativa de Rotação. (a) Valores de frequência da imagem de referência para $0.1\rho < r < \rho$. (b) Valores de frequência da imagem rotacionada para $0.1\rho < r < \rho$. (c) Valor médio em função do ângulo $h(\alpha)$ para $|G_0(\mathbf{u})|$ e $|G_1(\mathbf{u})|$. Fonte: Vandewalle (2006).

análise de correlação unidimensional. O exato ângulo de rotação pode ser calculado como o valor em que a correlação atinge o seu máximo.

Uma vez que a rotação é estabelecida, o deslocamento translacional entre as imagens é expresso no domínio de Fourier por um deslocamento de fase linear. A estimativa é calculada através do declive da diferença de fase entre as imagens. Para tornar a solução menos sensível a ruído, o cálculo utiliza o método de mínimos quadrados. No entanto quando as imagens em questão são de baixa resolução o efeito de *aliasing* modifica o conteúdo de frequência tornando o método impreciso. Portanto o deslocamento translacional passa a ser expresso como

$$G_1(\mathbf{u}) = \sum_{m=-M}^M e^{j2\pi(\mathbf{u}-m\mathbf{u}_s)^T t} G_0(\mathbf{u} - m\mathbf{u}_s), \quad (2.17)$$

onde t são os parâmetros de translação, \mathbf{u}_s a frequência de amostragem, e $2M + 1$ as cópias do espectro sobrepostas na frequência \mathbf{u} . O efeito de *aliasing* causa distúrbios na relação de fase linear entre $G_0(\mathbf{u})$ e $G_1(\mathbf{u})$. Porém, com o *aliasing* limitado, é possível utilizar os métodos citados acima, considerando apenas as frequências livres, ou pouco afetadas, pelo *aliasing*.

2.3 Métodos de Reconstrução

2.3.1 Iterative Back Projection - IBP

No método de reconstrução formulado por Irani (1991) a imagem de AR é estimada por um processo de retroprojeção iterativa (*iterative back-projection* - IBP) da diferença (o erro) entre simulações de BR provenientes da imagem de AR estimada na iteração anterior e as imagens de BR originais.

O método inicia com uma hipótese $\mathbf{f}^{(0)}$ da imagem de AR \mathbf{f} a ser obtida. A partir de $\mathbf{f}^{(0)}$, o sistema de captura da imagem é simulado para gerar um conjunto de imagens de BR $\mathbf{g}_k^{(0)}$ representando a correspondência com o conjunto original de imagens de BR \mathbf{g}_k . Se $\mathbf{f}^{(0)}$ representar corretamente a imagem de AR, então as imagens simuladas $\mathbf{g}_k^{(0)}$ devem ser idênticas as imagens de BR de entrada. Portanto, o erro encontrado no cálculo da imagem que contém a diferença entre \mathbf{g}_k e $\mathbf{g}_k^{(0)}$ é utilizado para aprimorar $\mathbf{f}^{(0)}$ através da retroprojeção do erro em $\mathbf{f}^{(0)}$ (o método é ilustrado na Figura 2.5). Esse processo é repetido iterativamente com o objetivo de minimizar a função de erro definida por

$$e^{(n)} = \sqrt{\sum_k \sum_{(x,y)} (\mathbf{g}_k(x,y) - \mathbf{g}_k^{(n)}(x,y))^2}. \quad (2.18)$$

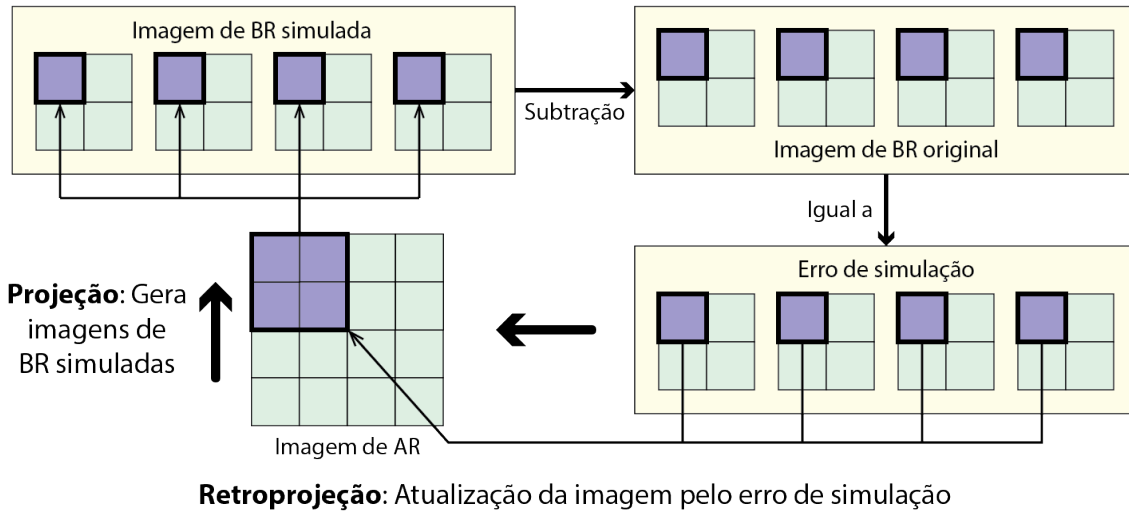


Figura 2.5: Diagrama semântico do algoritmo de super resolução do IBP (Adaptado de Park (2003)).

Considerando que o processo de captura das imagens pode ser expresso por

$$\mathbf{g}^{(n)}(\mathbf{y}) = \sum_{\mathbf{x}} \mathbf{f}^{(n)}(\mathbf{x}) h^{PSF}(\mathbf{x} - \mathbf{z}_y), \quad (2.19)$$

onde \mathbf{y} é um píxel de BR influenciado por \mathbf{x} um píxel de alta resolução, $\mathbf{f}^{(n)}$ é a aproximação de \mathbf{f} na iteração n , h^{PSF} é a função de propagação do ponto (*point spread function*) do sistema

de captura, e \mathbf{z}_y é o centro do campo receptivo \mathbf{y} em $\mathbf{f}^{(n)}$, o processo iterativo de refinamento para estimar a imagem de AR \mathbf{f} é expresso por

$$\mathbf{f}^{(n+1)}(\mathbf{x}) = \mathbf{f}^{(n)}(\mathbf{x}) + \sum_{\mathbf{y} \in \cup_k \mathbf{Y}_{k,\mathbf{x}}} (\mathbf{g}_k(\mathbf{y}) - \mathbf{g}_k^{(n)}(\mathbf{y})) \frac{(h_{\mathbf{xy}}^{BP})^2}{c \sum_{\mathbf{y}' \in \cup_k \mathbf{Y}_{k,\mathbf{x}}} h_{\mathbf{xy}'}^{BP}}, \quad (2.20)$$

onde $\mathbf{y}_{k,\mathbf{x}}$ denota o conjunto $\{\mathbf{y} \in \mathbf{g}_k \mid \mathbf{y} \text{ é influenciado por } \mathbf{x}\}$, c (constante) é um fator de normalização, e a função que simula h^{PSF} é representada por $h_{\mathbf{xy}}^{BP} = h^{BP}(\mathbf{x} - \mathbf{z}_y)$.

A atualização dos píxeis de AR \mathbf{x} em $\mathbf{f}^{(n)}$ se baseia em todos os píxeis de BR \mathbf{y} pelos quais \mathbf{x} influencia. A contribuição de um píxel de BR \mathbf{y} de uma imagem de entrada \mathbf{g}_k é o erro $(\mathbf{g}_k(\mathbf{y}) - \mathbf{g}_k^{(n)}(\mathbf{y}))$ multiplicado pelo fator $h_{\mathbf{xy}}^{BP}/c$. Portanto, píxeis de BR que são fortemente influenciados, também fortemente influenciam $\mathbf{f}^{(n+1)}(\mathbf{x})$, enquanto píxeis que são fracamente influenciados dificilmente influenciam $\mathbf{f}^{(n+1)}(\mathbf{x})$. Com a sobreposição dos campos receptivos de diferentes píxeis de BR, o novo valor de $\mathbf{f}^{(n+1)}(\mathbf{x})$ é influenciado por vários píxeis de BR. Todas as correções geradas por esses píxeis são combinados tomando a sua média ponderadas utilizando os coeficientes de h^{BP} como pesos.

Em alguns casos é possível que diferentes imagens de AR resultem em uma mesma imagem de BR, após o processo de captura. Por isso, existem múltiplas soluções para o problema, onde o algoritmo pode convergir para uma ou oscilar entre possíveis soluções. A escolha da hipótese inicial irá influenciar na convergência do algoritmo, porém não afeta a performance.

Outro fator determinante é a escolha de h^{BP} . Diferentemente de h^{PSF} , que representa as características do sensor, h^{BP} pode ser escolhida arbitrariamente. A convergência do algoritmo pode ser alcançada com diferentes h^{BP} , porém esta afeta diretamente as características da solução alcançada. Portanto, h^{BP} pode ser escolhida como uma restrição adicional para produzir uma solução final com especificidades diferentes.

2.3.2 Total Variation - TV

O segundo método se baseia na aplicação da “Variação Total”, ou *Total Variation* (TV) proposta por Rudin (1994) como termo de regularização no método de reconstrução descrito anteriormente. A regularização consiste na minimização de um problema envolvendo a oscilação da imagem. A quantidade total de mudança da imagem é penalizada pela norma da magnitude do gradiente de intensidade da imagem. A principal diferença entre a TV e os modelos de restauração de imagens baseados em mínimos quadrados é que a TV minimiza a norma L_1 do gradiente ao invés da norma L_2 (Chan, 2001). A tendência da aplicação da TV é uma imagem mais nítida, preservando as bordas da imagem.

Considerando uma imagem \mathbf{f} definida em um domínio contínuo $\Omega \subset \mathbb{R}^2$, denotada por $\mathbf{f}(\mathbf{x})$ para $\mathbf{x} = (x, y) \in \mathbb{R}^2$, e $v(\mathbf{x})$ uma versão de $\mathbf{f}(\mathbf{x})$ contaminada por ruído definida por

$$v(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + r(\mathbf{x}), \quad (2.21)$$

onde $r(\mathbf{x})$ representa um ruído aleatório com média 0 e variância σ^2 , o problema restrito que minimiza a variação total pode ser definido por

$$TV(\mathbf{f}) = \int_{\Omega} |\nabla \mathbf{f}| dx = \int_{\Omega} \sqrt{\mathbf{f}_x^2 + \mathbf{f}_y^2} dx dy, \quad (2.22)$$

onde $\nabla \mathbf{f}$ representa o gradiente, e dx o elemento de área de Ω . O problema possui as seguintes restrições relativas a média e variância, respectivamente, baseando-se nas suposições de $r(\mathbf{x})$, definidas por

$$\int_{\Omega} \mathbf{f} \, dx = \int_{\Omega} v \, dx, \quad (2.23)$$

$$\frac{1}{|\Omega|} \int_{\Omega} (\mathbf{f} - v)^2 \, dx = \sigma^2, \quad (2.24)$$

onde $|\Omega|$ é a área do domínio Ω da imagem. O problema anterior pode ser expresso como um problema irrestrito de otimização ponderando (2.22) por (2.24) e introduzindo o multiplicador Lagrangiano λ (uma vez que (2.23) é automaticamente satisfeita (Rudin, 1994))

$$J(\mathbf{f}) = \int_{\Omega} |\nabla \mathbf{f}| \, dx + \frac{\lambda}{2} \int_{\Omega} (\mathbf{f} - v)^2 \, dx. \quad (2.25)$$

O parâmetro λ controla o *trade-off* entre a regularidade e a fidelidade dos termos. A medida que o valor de λ diminui o peso do termo de regularidade aumenta. Com isso, λ se relaciona ao grau de filtragem da solução do problema de minimização (Buades, 2005).

A equação de Euler-Lagrange aplicada a J é definida por

$$-\nabla \cdot \left(\frac{\nabla \mathbf{f}}{|\nabla \mathbf{f}|} \right) + \lambda(\mathbf{f} - v) = 0. \quad (2.26)$$

Aplicando o método do gradiente descendente obtém-se

$$\frac{\partial \mathbf{f}}{\partial t} = \nabla \cdot \left(\frac{\nabla \mathbf{f}}{|\nabla \mathbf{f}|} \right) + \lambda(\mathbf{f} - v). \quad (2.27)$$

Caso $\lambda = 0$, ou não exista restrição apropriada define-se

$$\frac{\partial \mathbf{f}}{\partial t} = \nabla \cdot \left(\frac{\nabla \mathbf{f}}{|\nabla \mathbf{f}|} \right). \quad (2.28)$$

A TV tem sido utilizada com sucesso em problemas inversos considerados altamente mal condicionados, como o “*deblurring*” de imagens. Com a Variação Total as bordas retas tendem-se a se manter. Porém, os detalhes e textura podem ser super suavizados caso o valor de λ seja muito baixo (Buades, 2005).

2.3.3 Robust Super-Resolution - RSR

A proposta de Zomet (2001) introduz uma mudança no estimador (processo de otimização do método) da SR, aplicando a mediana do erro no critério de avaliação do resultado, ao invés da aplicação padrão que se utiliza da média. A abordagem resulta em um estimador robusto, contribuindo para a redução do erro causado pelas estimativas inconsistentes do processo de captura das imagens de BR. A utilização da mediana, além de produzir um aumento na robustez, é de baixo custo computacional e a perda em acurácia é imperceptível.

Considerando um conjunto de n imagens de entrada $\mathbf{g}_1, \dots, \mathbf{g}_n$ e o processo de captura definido pela Eq. (1.1), o erro quadrático total da reamostragem da imagem de AR \mathbf{f} se dá por

$$L(\mathbf{f}) = \frac{1}{2} \sum_{k=1}^n \|\mathbf{g}_k - D_k B_k M_k \mathbf{f}\|_2^2, \quad (2.29)$$

observando a derivada de L relacionada a \mathbf{f} , o gradiente de L é a soma dos gradientes calculados sobre as imagens de entrada

$$\begin{aligned} \mathbf{C}_k &= M_k^T B_k^T D_k^T (D_k B_k M_k \mathbf{f} - \mathbf{g}_k) \\ \nabla L(\mathbf{f}) &= \sum_{k=1}^n \mathbf{C}_k. \end{aligned} \quad (2.30)$$

O método de minimização iterativo baseado no gradiente mais simples atualiza a cada iteração a solução estimada, por

$$\mathbf{f}^{n+1} = \mathbf{f}^n + \lambda \nabla L(\mathbf{f}), \quad (2.31)$$

onde λ é um fator que define o dimensão do passo na direção do gradiente. No domínio da imagem, esta é uma versão do IBP. A cada iteração, a imagem de AR estimada é atualizada pelas imagens de entrada. A diferença entre as imagens de entrada e a imagem atualizada é retroprojetada na imagem de AR atual. Cada termo do somatório de \mathbf{C}_k na Eq. (2.30) corresponde a imagem da diferença retroprojeção.

Com o objetivo de introduzir a robustez ao processo, o somatório das imagens na Eq. (2.30) é substituída pela mediana escalar píxel-a-píxel, definida por

$$\nabla L(\mathbf{f})(x, y) \approx n \cdot \text{median}\{\mathbf{C}_k(x, y)\}_{k=1}^n. \quad (2.32)$$

Para obtenção de uma distribuição simétrica, a mediana pode aproximar a média com boa precisão, com uma quantidade suficiente de amostras. Na presença de ruídos a mediana apresenta uma maior robustez em relação à média. A estimativa pela mediana pode ser tendenciosa quando o ruído é organizado de forma assimétrica em relação à média. Neste caso um processo de detecção dessa tendência é aplicado baseado nas medições do ruído.

2.3.4 Normalized Convolution - NC

A contribuição proposta por Pham (2006a) está centralizada na aplicação da convolução normalizada, ou *Normalized Convolution* (NC), como termo de regularização no processo de reconstrução da imagem de AR. O sinal local é aproximado através de uma projeção em um subespaço gerado por um conjunto de funções de base. A aplicação do método garante melhorias na taxa sinal-ruído e minimiza as influências sofridas pelos erros ocorridos na fase de registro.

A Convolução Normalizada, ou *Normalized Convolution* (NC) é uma técnica de modelagem de sinal local a partir de projeções em um conjunto de funções de base. Embora vários tipos de base podem ser utilizados, bases polinomiais são geralmente utilizadas, que podem ser definidas por: $\{\mathbf{l}, \mathbf{x}, \mathbf{y}, \mathbf{x}^2, \mathbf{y}^2, \mathbf{xy}\}$, onde $\mathbf{l} = [1 \ 1 \ \dots \ 1]^T$ (N entradas), $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$, $\mathbf{x}^2 = [x_1^2 \ x_2^2 \ \dots \ x_n^2]^T$, e assim por diante é construída a partir de coordenadas locais por N amostras

de entrada. A utilização de funções de base polinomiais faz com que a NC tradicional seja equivalente a expansão local da série de Taylor. Considerando uma vizinhança local centralizada em $\mathbf{s}_0 = \{x_0, y_0\}$, o valor de intensidade na posição $\mathbf{s} = \{x + x_0, y + y_0\}$ é aproximado por uma expansão polinomial definida por

$$\hat{\mathbf{f}}(\mathbf{s}, \mathbf{s}_0) = p_0(\mathbf{s}_0) + p_1(\mathbf{s}_0)x + p_2(\mathbf{s}_0)y + p_3(\mathbf{s}_0)x^2 + p_4(\mathbf{s}_0)xy + p_5(\mathbf{s}_0)y^2 + \dots, \quad (2.33)$$

onde $\{x, y\}$ são as coordenadas locais da amostra \mathbf{s} relacionada ao centro de análise, e $\mathbf{s}_0 \cdot \mathbf{p}(\mathbf{s}_0) = [p_0 \ p_1 \ \dots \ p_m]^T(\mathbf{s}_0)$ são os coeficientes de projeção nas funções de base polinomiais correspondentes em \mathbf{s}_0 .

A escolha da ordem polinomial depende da aplicação. Portando há um *trade-off* entre a velocidade do método e a complexidade do modelo a ser aplicado. Quanto maior a complexidade das estruturas a serem modeladas, maior é a ordem da NC necessária, gerando um processo mais custoso.

Para calcular os coeficientes de projeção \mathbf{p} a uma posição de saída \mathbf{s}_0 , a aproximação do erro é minimizada através da extensão de uma função a centralizada em \mathbf{s}_0

$$\varepsilon(\mathbf{s}_0) = \int (\mathbf{f}(\mathbf{s}) - \hat{\mathbf{f}}(\mathbf{s}, \mathbf{s}_0))^2 c(\mathbf{s}) a(\mathbf{s}, \mathbf{s}_0) d\mathbf{s}, \quad (2.34)$$

onde o sinal de certeza $0 \leq c(\mathbf{s}) \leq 1$ especifica a confiabilidade da medição em \mathbf{s} , com o zero representando dados completamente não confiáveis e um representando dados muito confiáveis. Ambos c e a agem como pesos para o cálculo do erro quadrático. A solução pela regressão por mínimos na forma de matriz pode ser definida por

$$\mathbf{p} = (\mathbf{B}^T \mathbf{W} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{W} \mathbf{F}, \quad (2.35)$$

onde \mathbf{F} é uma matriz com as intensidades de $\mathbf{f}(\mathbf{s})$, $\mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_m]$ é a matriz que contém as funções de base, e $\mathbf{W} = \text{diag}(\mathbf{c}) \cdot \text{diag}(\mathbf{a})$ é a matriz diagonal construída pelo produto do sinal de certeza \mathbf{c} pela aplicabilidade amostrada \mathbf{a} .

2.3.5 G-PMSR

A abordagem proposta por Maiseli (2014) possui um termo de regularização com um expoente variável adaptativo $\gamma(z)$ em função da magnitude do gradiente. A abordagem assumida pelo método de reconstrução é local. Portanto, o expoente assume diferentes valores ao percorrer a imagem. Quando a magnitude do gradiente é alta (indicativo de regiões contendo bordas) o expoente assume valores elevados, e quando o gradiente apresenta baixa magnitude (indicativo de regiões homogêneas) o expoente tende a zero. O objetivo é que o método favoreça a preservação das bordas, bem como a redução do ruído em regiões homogêneas.

O termo de regularização proposto pode ser definido por

$$0 = \beta_1 \nabla \cdot \left(\frac{\nabla \mathbf{f}}{(1 + [|\nabla \mathbf{f}| \lambda]^{\gamma(z)})} \right) - \beta_2 (\mathbf{f} - \mathbf{f}_0), \quad (2.36)$$

onde \mathbf{f} representa a imagem de AR desejada, β_1 é um parâmetro de regularização, β_2 é um limiar de fidelidade e $0 \leq \gamma(z) \leq 2$.

Combinando a solução do problema de reconstrução robusto (2.30) com o termo de regularização proposto resulta em um sistema iterativo descrito por

$$\frac{\partial \mathbf{f}}{\partial t} = \sum_{k=1}^K (M_k^T B_k^T D_k^T (D_k B_k M_k \mathbf{f} - \mathbf{g}_k)) + \beta_1 K \nabla \cdot \left(\frac{\nabla \mathbf{f}}{(1 + [|\nabla \mathbf{f}|/\lambda]^{\gamma(z)})} \right) - \beta_2 K (\mathbf{f} - \mathbf{f}_0). \quad (2.37)$$

O componente $\gamma(z)$ é variável e se atualiza de forma automática ajustando a regularização com base nas estruturas locais da imagem. Sendo

$$Q_{SR}(\mathbf{f}) = \sum_{k=1}^K (M_k^T B_k^T D_k^T (D_k B_k M_k \mathbf{f} - \mathbf{g}_k)) \quad (2.38)$$

$$R(\mathbf{f}) = \beta_2 K (\mathbf{f}_0 - \mathbf{f}),$$

é possível perceber que quanto $\gamma(z) = 0$, a equação se define por

$$\frac{\partial \mathbf{f}}{\partial t} = Q_{SR}(\cdot) + \beta_1 K \Delta \mathbf{f} + R(\cdot), \quad (2.39)$$

onde nesse caso a regularização favorece as partes homogêneas da imagem. Quando $\gamma(z) = 1$ a equação se define por

$$\frac{\partial \mathbf{f}}{\partial t} = Q_{SR}(\cdot) + \beta_1 K \nabla \cdot \left(\frac{\nabla \mathbf{f}}{(1 + [|\nabla \mathbf{f}|/\lambda])} \right) + R(\cdot), \quad (2.40)$$

onde a regularização se aproxima à proposta por Charbonnier (1994). E em um terceiro caso onde $\gamma(z) = 2$ equação se define por

$$\frac{\partial \mathbf{f}}{\partial t} = Q_{SR}(\cdot) + \beta_1 K \nabla \cdot \left(\frac{\nabla \mathbf{f}}{(1 + [|\nabla \mathbf{f}|/\lambda]^2)} \right) + R(\cdot), \quad (2.41)$$

onde a esse é o modelo base para a abordagem proposta em Perona (1994). Com o objetivo de tornar o termo $\gamma(z)$ auto ajustável durante o processo, a seguinte equação é proposta

$$\gamma(z) = 2(1 - \kappa \exp(-|G_\sigma * \mathbf{f}|)), \quad (2.42)$$

onde $\kappa \in [0, 1]$ é o parâmetro de refinamento e G_σ é um *kernel* Gaussiano.

Capítulo 3

Experimentos

Este capítulo descreve os experimentos realizados nesse trabalho divididos em duas etapas. A primeira destaca os testes realizados com imagens de BR simuladas. Essa etapa tem como principal objetivo avaliar a melhor combinação entre técnicas de registro e reconstrução para o problema de SR. Uma vez selecionada as melhores técnicas, dá-se início aos testes com dados de imagens de BR reais, onde será avaliada a eficácia da combinação escolhida para casos reais. Aqui está contida, para ambas etapas, a descrição dos dados utilizados, a motivação para os casos de teste propostos, a metodologia utilizada para avaliar as técnicas e os resultados obtidos.

3.1 Introdução

A avaliação de uma imagem de AR resultante da SR é muitas vezes qualitativa. Em conjuntos reais de imagens de BR, não se sabe com exatidão quais valores os píxeis reconstruídos devem assumir, portanto não é possível definir o resultado ótimo do problema. Duas etapas de experimentos são propostas no trabalho, uma com imagens de BR simuladas e a segunda com imagens de BR reais.

A primeira fase oferece um ambiente controlado para os experimentos. Com base em uma imagem de AR, um conjunto de imagens de BR é gerado. As versões de BR são geradas por meio da aplicação de diferentes transformações de escala, translação e rotação na imagem de AR. As técnicas de registro e reconstrução, utilizam-se dessas imagens construídas para gerar uma hipótese da imagem de AR. Por meio da comparação entre a hipótese gerada e a imagem de AR original, é possível avaliar quantitativamente os métodos investigados.

Na segunda fase os conjuntos de imagens de BR utilizados são provenientes de sequências de imagens de uma mesma cena com transformações desconhecidas, onde os métodos serão novamente avaliados qualitativamente, por não haver uma imagem base para comparação. Com esses experimentos espera-se obter as técnicas que se destaquem nas fases de registro e reconstrução e assim definir a melhor combinação para aplicação da SR.

3.2 Experimentos com imagens simuladas

3.2.1 Metodologia

A primeira etapa de experimentos constitui-se de experimentos com imagens de BR simuladas por meio de uma imagem de AR de referência. Primeiramente, 128 imagens foram selecionadas de diferentes bases de dados publicadas na literatura. Dentre essas estão contidas imagens aéreas, termográficas, de ressonância magnética e outras imagens de conteúdo aleatório.

Após a seleção das imagens base para os experimentos, diferentes conjuntos de imagens de BR foram gerados a partir dessas imagens. Esses conjuntos foram utilizados em 6 casos de testes para avaliar os métodos em condições distintas. Cada uma das 128 imagens submetidas aos 6 casos de teste resultou em 768 conjuntos de imagens de BR para avaliação dos métodos de registro e reconstrução.

Os métodos de registro são avaliados da seguinte forma. Uma vez que um método estima as transformações de translação e rotação entre as imagens BR de um conjunto, esses valores são comparados com aqueles utilizados para gerar o conjunto de imagens de BR. Essa comparação é feita com base no erro de estimação, utilizando a Raiz do Erro Quadrático Médio (REQM). Os operadores avaliados são: a) translação na horizontal (Δx), b) translação na vertical (Δy) e c) rotação (θ), medida em graus. Os métodos de reconstrução são avaliados com base na imagem original utilizada para gerar as imagens de BR. Como a reconstrução é uma etapa posterior ao registro, os valores reais dos deslocamentos entre as imagens são fornecidos aos métodos. Uma vez gerada uma hipótese de AR essa é comparada com os parâmetros utilizados na construção das imagens de BR, utilizando a Relação Sinal/Ruído de Pico (RSRP), o Coeficiente de Correlação de Pearson (CCP) e o Tempo de execução (TE).

Ao final dessa etapa de avaliação será possível discernir entre os métodos avaliados com aqueles que se destacam nas etapas de registro e reconstrução. A motivação é obter a melhor combinação de técnicas para aplicação da SR.

Para realizar os experimentos das duas etapas utilizou-se uma plataforma computacional contendo, um processador Intel(R) Core (TM) i7-4510U 2.00 GHz, 8 GB de memória RAM, usando um sistema operacional Windows 8.1 64-bit e o software MATLAB (2014). Os códigos dos métodos de registro e reconstrução das imagens foram obtidos através do software, baseado em MATLAB, disponibilizado pelo LCAV (2015).

3.2.2 Métricas de avaliação

Raiz do Erro Quadrático Médio

A Raiz do Erro Quadrático Médio (REQM, ou *Root Mean Square Error*) tem sido amplamente usada como uma métrica estatística padrão para mensurar a performance de um modelo (Chai, 2014). Nesse trabalho ela é utilizada como medida de avaliação de erro para os métodos de registro.

A REQM compara os valores obtidos como resultado da estimação de transformação entre as imagens e os valores utilizados para formar as imagens de BR simuladas - a REQM também é utilizada para avaliar o registro em Robinson (2004), Pham (2006a) e Maiseli (2014). Uma

vez que o objetivo é a minimização do erro, os métodos que apresentam uma menor REQM possuem melhores resultados. A REQM pode ser definida como

$$REQM = \sqrt{EQM}, \quad (3.1)$$

$$EQM = \frac{\sum_{i=1}^n (O_i - A_i)^2}{n}, \quad (3.2)$$

onde n representa o número de amostras (número de imagens a ser registrada, ou a quantidade de píxeis na imagem de AR) do conjunto, O_i é o valor de referência e A_i é o valor estimado.

Tempo de Execução

O tempo de execução também é uma métrica utilizada para as duas classes de métodos avaliados. Esse critério tem como objetivo avaliar quão custoso é o método ao realizar a estimativa. O TE é avaliado como o tempo em segundos que o método consome para estimar o registro entre as imagens de BR ou a reconstrução da hipótese de AR. Na contagem do TE não está inserida o tempo para carregar as imagens no algoritmo, uma vez que esse processo é o mesmo para todos os métodos.

Relação Sinal/Ruído de Pico

A Relação Sinal/Ruído de Pico (RSRP, ou *Peak Signal-to-Noise Ratio*) é conhecida como uma das métricas mais simples medir a similaridade entre imagens e é frequentemente utilizada no âmbito de processamento de imagens - métodos como Babacan (2011) e Maiseli (2014) utilizam a RSRP para avaliar o resultado da aplicação de um processo de SR. A RSRP é definida como a razão entre o valor máximo possível de um sinal e a influência do ruído que corrompe a fidelidade de representação do sinal (Zhu, 2013). Relações que apresentam um maior valor de RSRP expressam que as imagens possuem maior semelhança entre si. A RSRP pode ser definida como

$$RSRP = 10 \log_{10} \frac{255^2}{EQM} \quad (3.3)$$

Coefficiente de Correlação de Pearson

Outra métrica para comparar o resultado dos métodos de reconstrução com a imagem de AR de referência é o Coeficiente de Correlação de Pearson (CCP). Adler (2010) afirma que o CCP é uma métrica útil para medir a correlação entre imagens. Imagens com índice elevado de correlação apresentam uma maior semelhança entre si. O CCP em imagens é medido na escala de $[-1, 1]$ onde -1 representa que não existe nenhuma correlação entre as imagens e 1 estabelece que a correlação entre as imagens é perfeita. O CCP pode ser definido como

$$CCP = \frac{\sum_m \sum_n (O_{mn} - \bar{O})(A_{mn} - \bar{A})}{\sqrt{(\sum_m \sum_n (O_{mn} - \bar{O})^2) (\sum_m \sum_n (A_{mn} - \bar{A})^2)}} \quad (3.4)$$

onde O e A são as imagens a serem comparadas, $\bar{O} = \text{média}(O)$, $\bar{A} = \text{média}(A)$ e, m e n se referem as posições dos píxeis nas imagens.

3.2.3 Dados

Um conjunto de 128 imagens foram selecionadas para avaliação na primeira etapa de experimento. A Figura 3.1 ilustra 15 exemplos das diferentes imagens utilizadas no conjunto de teste. A base de imagens possui: 1 imagem sintética (adaptada do gráfico de resolução (EIA 1956)); 39 imagens aéreas e 28 imagens (Weber, 1997); 25 imagens termográficas (Morris, 2007); 35 imagens de ressonância magnética (NZBRI, 2015).



Figura 3.1: Exemplos de imagens utilizadas na primeira fase se experimentos.

A seleção das imagens tem como objetivo envolver ambientes favoráveis a aplicação da SR. Nas aplicações envolvendo imagens aéreas é possível obter imagens de AR. No entanto, regiões de interesse selecionadas nas imagens podem apresentar BR. A distância entre o sensor e a cena pode provocar que objetos relativamente grandes na superfície, sejam resumidos em escalas de píxeis. Portanto, a aplicação da SR pode ser interessante para melhor análise da região de interesse em questão. As imagens térmicas e de ressonância magnética geralmente são de BR. Os elevados custos envolvidos na fabricação de sensores AR, muitas vezes, inviabiliza a sua fabricação e utilização para essas aplicações.

As imagens foram selecionadas de bases de dados com características diferentes. Visando padronizar as imagens para os experimentos, todas foram convertidas para a escala de cinza em 8 bits e com resolução de 480x480 píxeis. Conjuntos de imagens de BR serão gerados a partir

de cada uma das imagens selecionadas, de acordo com os casos de teste propostos.

3.2.4 Casos de Teste

Na construção dos casos de teste para avaliação dos métodos de registro e reconstrução, foram considerados três aspectos: fator de escala, relação entre fator de escala versus quantidade de imagens no conjunto de BR e a distribuição das imagens BR no plano de AR.

As imagens selecionadas foram utilizadas para formar conjuntos de imagens de BR para avaliação dos métodos de registro e reconstrução. O processo para construir esses conjuntos é exemplificado na Figura 3.2. O objetivo é simular o processo de um sistema de captura de imagem ao registrar amostras de uma mesma cena de AR no domínio contínuo. O conjunto de imagens a ser construído possui versões discretas de BR da imagem de AR, apresentando pequenas diferenças de deslocamento entre si. Para representar as degradações na imagem durante a captura, primeiramente filtros de *blur* e ruído são aplicados na imagem de referência (a), resultando em uma imagem de AR degradada (b). Uma vez concluído esse processo, a imagem é reduzida por um FE, resultando em uma imagem de BR de referência (c). Esse fator representa a taxa de redução de resolução entre a imagem de AR e as imagens de BR. Uma vez determinada a quantidade de imagens do conjunto de teste, diferentes combinações de transformações de translação e rotação são atribuídas a (c) gerando as diferentes versões de BR (d) da imagem de AR, ou seja, as imagens base para o processo de SR.

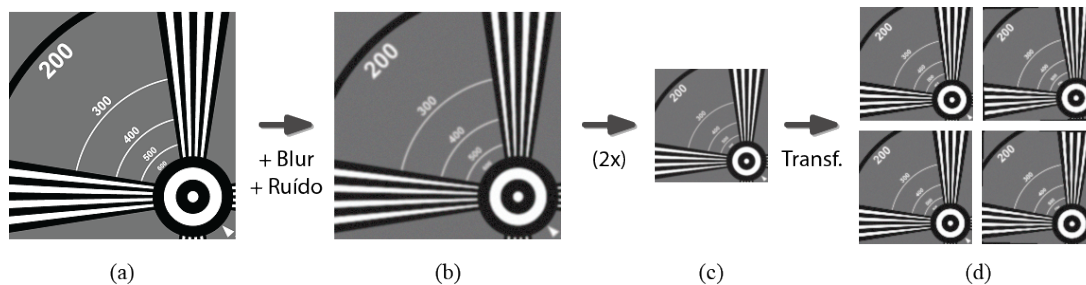


Figura 3.2: Processo para gerar o conjunto de imagens de BR a partir de uma imagem de AR. (a) Imagem de AR; (b) Imagem apresentando degradações de ruído e blur gaussiano; (c) Imagem reduzida por um FE= 2; (d) Versões de BR da imagem com diferentes deslocamentos de sub-píxel.

As combinações de transformação definidas não implicam que todas as imagens possuam transformações de translação e rotação. É necessário que a combinação envolvendo uma translação e/ou rotação de uma imagem seja distinta das outras imagens do conjunto. Por exemplo, duas imagens podem apresentar os mesmos valores de translação relativo a Figura 3.2c, mas com valores de rotação diferentes. Portanto, as imagens de BR podem apresentar transformações de: translação, rotação ou translação/rotação. Para que as imagens do conjunto sejam caracterizadas como diferentes versões de uma mesma cena de AR, os valores dos operadores de translação possuem escala de sub-píxel.

O FE é a taxa pela qual uma imagem tem a sua resolução aumentada ou diminuída. No caso da SR, o que deve ser avaliado é, o quanto a cena, representada pelas imagens de baixa

resolução, pode ter a sua resolução aumentada com resultados satisfatórios? Como apresentado por Yang (2014) métodos de SR com apenas uma imagem apresentam restrições a fatores de escala maiores que 4. Portanto, três valores de FE foram escolhidos: 2, 4 e 6. Esses valores foram determinados por apresentar níveis de complexidade diferentes. Sendo 2 um ambiente de menor complexidade e 6 o de maior complexidade.

O segundo aspecto mostra a relação entre a quantidade de informação que é disponibilizada no problema e o FE escolhido. O exemplo descrito na Figura 3.3 apresenta um conjunto de imagens de BR com características ideais para formar uma imagem de AR. A partir da Figura 3.3a, aplicando diferentes operadores de translação e reduzindo a sua resolução com um FE=4, forma-se um conjunto de FE² imagens de BR Figura 3.3b. Neste caso específico os valores de deslocamento translacional aplicados a cada uma das imagens estão uniformemente espalhados no plano de AR. Com essa quantidade de imagens de BR, que não apresentam

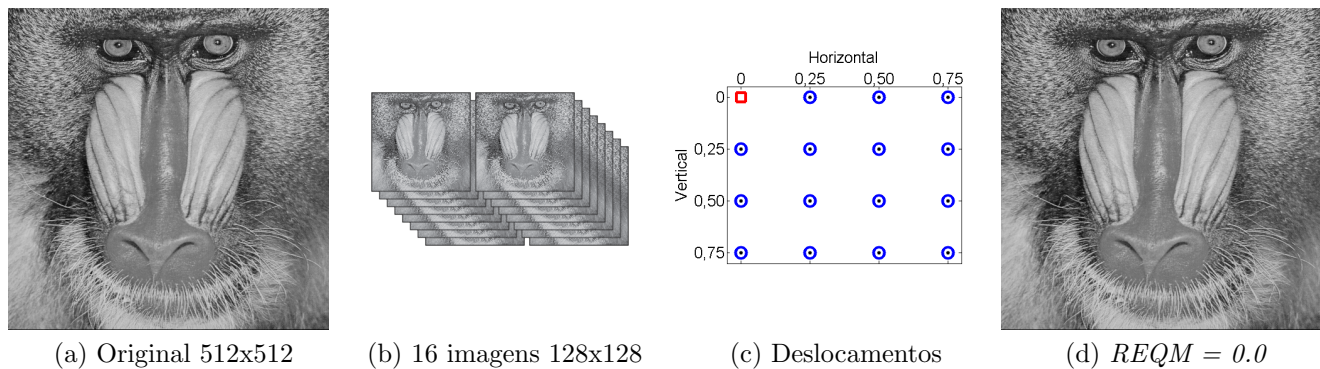


Figura 3.3: Representação de um caso ideal para fusão de imagens de BR, sem processo de degradação, em uma imagem de AR, onde a quantidade de imagens é igual ao FE² e os deslocamentos das imagens de BR são bem distribuídos no plano de AR. Uma imagem de AR (a) gera um conjunto de 16 imagens de BR (b) com FE igual a 4, sem degradações na imagem e deslocamentos uniformemente distribuídos no plano de AR (c), gerando uma hipótese da imagem de AR idêntica a original, apresentando $REQM = 0$.

degradações, e deslocamentos distribuídos uniformemente no plano de AR, é possível recuperar perfeitamente a imagem de AR Figura 3.3d. Isso é possível pois a quantidade de informação, ou seja, a quantidade de píxeis presente na imagem original é igual ao número de píxeis do conjunto de imagens de BR, 262.144 píxeis. Uma vez que deslocamentos entre as imagens de BR permite que estas sejam uniformemente distribuídas no plano de AR, os píxeis de BR preenchem perfeitamente a imagem de AR.

No entanto, existem casos onde a quantidade de imagens de BR é baixa em relação ao FE escolhido. Esse caso é exemplificado na figura Figura 3.4. Neste caso o FE usado é o mesmo do exemplo anterior, porém a quantidade de imagens disponíveis é menor Figura 3.4b. É possível notar que o plano de AR não foi totalmente preenchido, mesmo que contando com uma distribuição uniforme das imagens no plano de AR (Figura 3.4c). Com isso a imagem de AR gerada, Figura 3.4d, apresenta maiores imperfeições comparada com a Figura 3.3d. Os casos de testes propostos também apresentarão cenários com quantidade de imagens de BR

diferentes. Em alguns casos de teste, a quantidade de imagens será igual ao FE^2 , nos outros casos a quantidade será igual a $FE^2/2$. Essa característica influencia a complexidade do caso de teste.

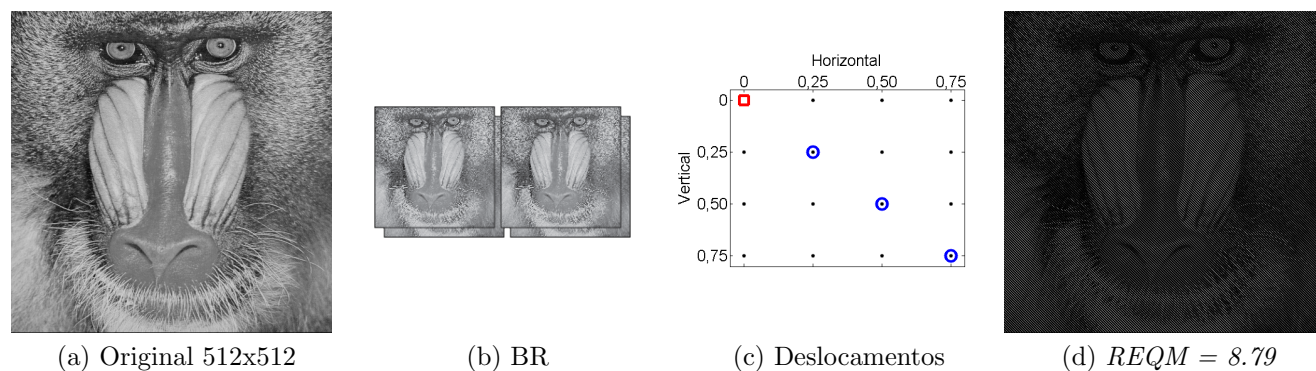


Figura 3.4: Representação de um caso para fusão de imagens de BR, sem processo de degradação, em uma imagem de AR, onde a quantidade de imagens é inferior ao FE^2 e os deslocamentos das imagens de BR são bem distribuídos no plano de AR. Uma imagem de AR (a) gera um conjunto de 4 imagens de BR (b) com FE igual a 4, sem degradações na imagem e deslocamentos com distribuição uniforme no plano de AR (c), gerando uma hipótese da imagem de AR apresentando $REQM = 8.79$.

Além da quantidade de imagens, a forma com que essas imagens estão distribuídas no plano de AR também é relevante ao problema de SR. A Figura 3.5 traz um exemplo desse caso. Nesse exemplo o $FE = 4$ e a quantidade de imagens = FE^2 . Entretanto, não existe uma distribuição adequada dessas imagens no plano de AR (Figura 3.5c). As diferenças de deslocamento entre algumas imagens são muito pequenas, implicando em algumas sobreposições de píxeis na imagem de AR. Mesmo que exista uma quantidade de informação considerada suficiente, parte dessa informação não é relevante ao problema. Assim, mais uma vez, a imagem de AR resultante possui imperfeições Figura 3.5d.

Considerando os aspectos citados, 6 casos de teste foram propostos para avaliar os métodos de registro e reconstrução de imagens de BR, descritos na Tabela 3.1. É possível notar na tabela que os casos de teste com numeração par possuem nível de complexidade maior, dado que para os três FE escolhidos os casos 2, 4 e 6 possuem uma quantidade inferior de imagens e também distribuição não uniforme. A resolução das imagens no conjunto de BR em cada teste está condicionada ao fator de escala, lembrando que as imagens de base estão padronizadas com resolução de 480x480 píxeis. A Tabela 3.2 exemplifica os operadores das transformações de translação e rotação para os casos de teste 3 e 4. O caso de teste 3 possui um FE igual a 4 e 16 imagens e é possível notar que os valores dos operadores são uniformemente distribuídos e com precisão de 2 casas decimais. O caso de teste 4 apresenta um nível de complexidade maior. Apesar de apresentar o mesmo FE a quantidade de imagens é a metade do caso de teste 3 e além disso a distribuição das imagens no plano de AR é aleatória com precisão de 15 casas decimais. O objetivo dos diferentes casos de teste é proporcionar ambientes com desafios diferentes para cada método de registro e reconstrução.

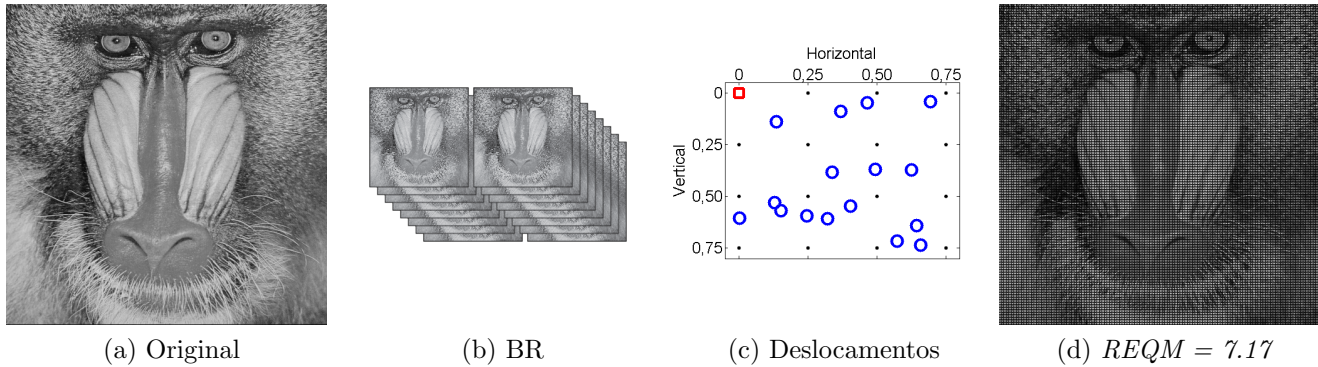


Figura 3.5: Representação de um caso para fusão de imagens de BR, sem processo de degradação, em uma imagem de AR, onde a quantidade de imagens é igual ao FE^2 e os deslocamentos das imagens de BR não são bem distribuídos no plano de AR. Uma imagem de AR (a) gera um conjunto de 16 imagens de BR (b) com FE igual a 4, sem degradações na imagem e deslocamentos aleatoriamente distribuídos no plano de AR (c), gerando uma hipótese da imagem de AR apresentando $REQM = 7.17$.

Tabela 3.1: Casos de teste propostos para a primeira etapa dos experimentos.

Caso	Imagens	Fator de Escala	Distribuição	Resolução
1	4	2	Uniforme	240x240
2	2	2	Aleatório	240x240
3	16	4	Uniforme	120x120
4	8	4	Aleatório	120x120
5	36	6	Uniforme	80x80
6	18	6	Aleatório	80x80

Tabela 3.2: Movimentos de translação (Δx , Δy) e rotação (θ em graus) em relação a imagem de referência para dezesseis imagens de BR no caso de teste 3 e para oito imagens de BR no caso de teste 4.

Caso	Tran	iRef	i2	i3	i4	i5	i6	i7	i8	i9	i10	i11	i12	i13	i14	i15	i16	
3	Δx	0.00	0.00	0.00	0.00	0.25	0.25	0.25	0.25	0.50	0.50	0.50	0.50	0.75	0.75	0.75	0.75	
	Δy	0.00	0.00	0.25	0.50	0.75	0.00	0.25	0.50	0.75	0.00	0.25	0.50	0.75	0.00	0.25	0.50	0.75
	θ	0.00	0.00	0.00	0.00	1.00	-1.00	1.00	-1.00	2.00	-2.00	2.00	-2.00	3.00	-3.00	3.00	-3.00	
4	Δx	0.00	0.02	0.14	0.03	0.70	0.57	0.58	0.13	-	-	-	-	-	-	-	-	
	Δy	0.00	0.06	0.13	0.31	0.60	0.55	0.09	0.70	-	-	-	-	-	-	-	-	
	θ	0.00	-1.49	2.11	0.82	-3.23	-1.76	-2.88	3.32	-	-	-	-	-	-	-	-	

3.2.5 Resultados

Parametrização dos métodos de registro

Dos quatro métodos avaliados apenas o Keren (1988) não possui parâmetros de entrada a serem ajustados. Os métodos de Marcel (1997) e Lucchese (2000) tem um parâmetro a ser ajustado, o fator de interpolação das TF das imagens de BR. Nesses métodos as imagens de BR passam por um processo de interpolação visando aumentar a precisão do registro. O método

de Vandewalle (2006) possui dois parâmetros. O primeiro é o valor do raio que será utilizado para estimar a transformação de rotação. O segundo é o número de componentes de baixa frequência utilizados para estimar as translações entre as imagens. Portanto, primeiramente, alguns experimentos foram realizados com o objetivo de definir os valores de parâmetro visando aumentar a precisão de cada métodos.

Com diferentes valores de parâmetro os métodos foram submetidos aos seis casos de teste descritos na Sessão 3.2.4 para as 128 imagens selecionadas. Para auxiliar na escolha dos parâmetros foram calculadas, para todos testes, a média da REQM do registro de rotação e translação bem como o TE. Para os métodos de Marcel (1997) e Lucchese (2000) foram testados fatores de interpolação com valores variando no seguinte intervalo $[1, 2, 3, \dots, 14, 15]$. No caso do método de Vandewalle (2006) os valores de raio avaliados variam no intervalo de, $[0.05r, 0.10r, 0.15r, \dots, 0.90r, 0.95r]$, onde r representa o raio da imagem. A quantidade de componentes de baixa frequência foram avaliados no intervalo de, $[1, 2, 3, \dots, 14, 15]$.

A Tabela 3.3 mostra os resultados para avaliação de fatores de interpolação para o método de Marcel (1997), com os seguintes critérios: média da REQM para os movimentos de horizontais de translação (Δx); média da REQM para os movimentos de verticais de translação (Δy); média entre Δx e Δy ; média da REQM para os movimentos de rotação (θ), e; média do TE - esses critérios também foram usados para o método de Lucchese (2000) e Vandewalle (2006). A Figura 3.6 traz os gráficos que mostram a variação de cada um dos critérios de avaliação por fator de interpolação.

Tabela 3.3: Resultado da avaliação do fator de interpolação das TF das imagens de BR para o método de Marcel (1997), apresentando: média da REQM para os movimentos horizontais de translação (Δx); média da REQM para os movimentos verticais de translação (Δy); média entre Δx e Δy ; média da REQM para os movimentos de rotação (θ), e; média do TE.

Critérios	Fator de Interpolação														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Δx	4.51	4.72	4.61	4.56	4.53	4.49	4.49	4.51	4.50	4.47	4.50	4.49	4.47	4.46	4.48
Δy	4.58	4.82	4.57	4.55	4.39	4.42	4.42	4.38	4.44	4.43	4.44	4.43	4.40	4.40	4.37
$med(\Delta x, \Delta y)$	4.55	4.77	4.59	4.56	4.46	4.45	4.45	4.45	4.47	4.45	4.47	4.46	4.43	4.43	4.43
θ	3.90	3.90	4.07	4.07	4.07	4.14	4.26	4.26	4.28	4.93	5.86	6.27	6.84	7.72	8.57
Tempo	0.10	0.20	0.39	0.62	0.90	1.25	1.57	1.94	2.43	2.66	3.03	3.38	3.54	3.74	3.93

Observando as Figuras 3.6a, 3.6b e 3.6c é possível notar uma disparidade entre os resultados obtidos. O erro de registro de translação tende a diminuir ao aumentar o valor do fator de interpolação, ao passo que, no registro de rotação e no TE, o comportamento é inverso. Portanto, se faz necessário a escolha de um fator que atenda os objetivos com um bom “*custo x benefício*”. A Figura 3.6d mostra o espaço de decisão relacionando as soluções da estimativa de translação e de rotação. Para auxiliar na decisão de escolher o fator de interpolação, o conceito de Dominância Pareto foi utilizado. Em um problema multiobjetivo de minimização, pode-se dizer que uma solução s_1 Pareto Domina (ou, simplesmente, domina) uma solução s_2 se em pelo menos um dos objetivos, $f(s_1) \leq f(s_2)$ e/ou $f(s_1) \neq f(s_2)$, é válido. Uma solução pode ser considerada indiferente ou não dominada quando não existe uma solução que a domine (Arroyo, 2002). A Figura 3.7 exemplifica a dominância Pareto de soluções.

Na Figura 3.6d foram relacionados os resultados de cada um dos fatores pela média da

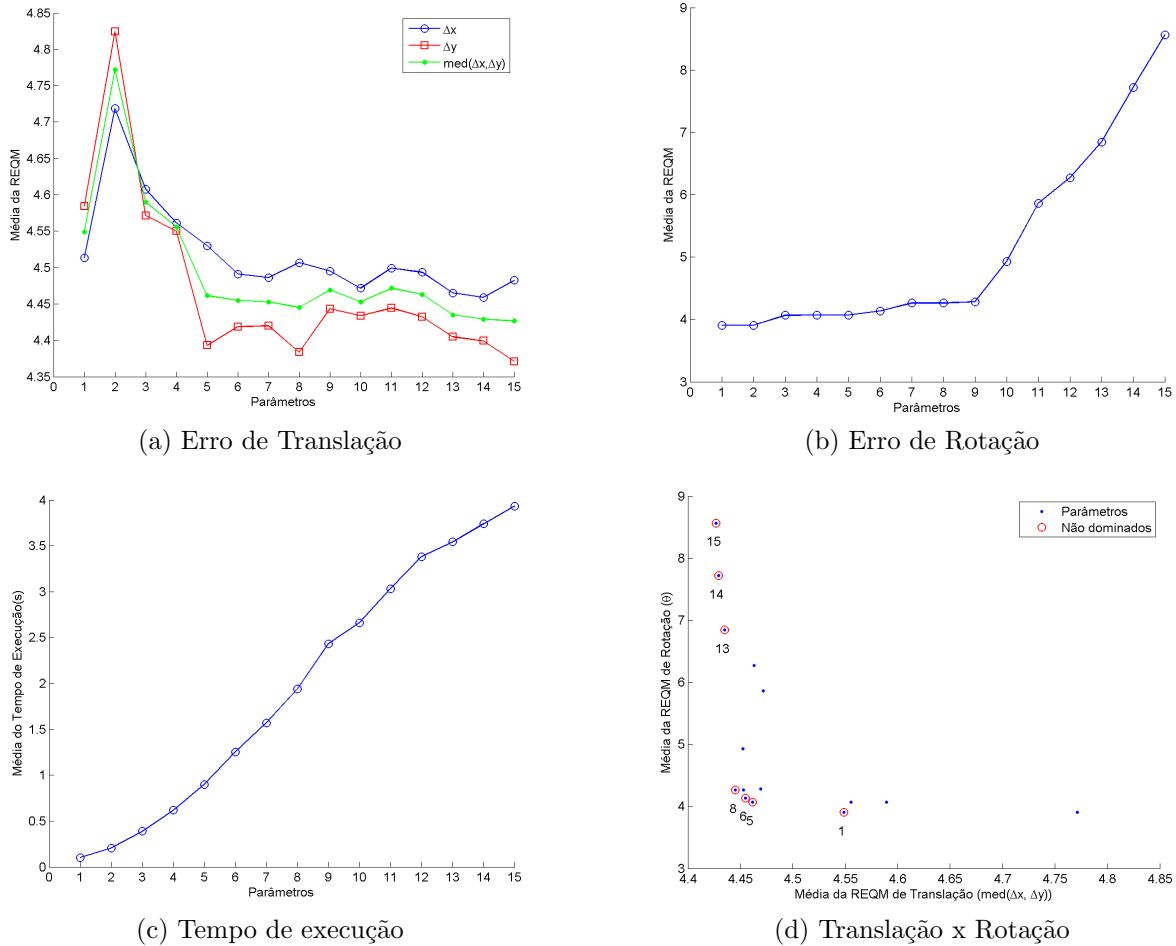


Figura 3.6: Resultados da avaliação dos fatores de interpolação aplicados ao método de Marcel (1997). a) Média da REQM para o registro de translação; b) Média da REQM para o registro de rotação; Média do TE em segundo; d) Relação entre os erros de registro para a translação e rotação.

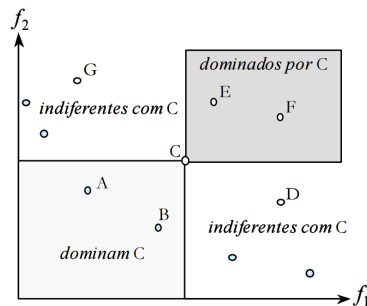


Figura 3.7: Dominância de Pareto no espaço de soluções. Adaptado de Arroyo (2002)

REQM de translação ($med(\Delta x, \Delta y)$) e da média da REQM de rotação (θ). Os pontos em destaque representam as soluções não dominadas (indiferentes) dentre os 15 fatores de interpo-

lação avaliados. Dentre esses, os que apresentam a melhor relação entre os dois objetivos são os valores, 5, 6 e 8. O tempo de execução não é considerado no espaço de decisão pois o seu comportamento se assemelha ao comportamento do erro de rotação. Comparando apenas a Figura 3.6b e Figura 3.6c, claramente o valor escolhido seria o 1. A relação entre os resultados do erro de rotação e tempo de execução não apresenta comportamento divergente. Após a pré-seleção dos valores 5, 6 e 8 o tempo de execução é analisado para auxiliar na tomada de decisão. Observando a Tabela 3.3 o fator 5 possui a menor média de TE, dentre os valores pré selecionados. Portanto, o valor 5 foi escolhido para o ser fator de interpolação das TF da imagens de BR a serem registradas para o método de Marcel (1997) nos experimentos.

A Tabela 3.4 e a Figura 3.8 mostram os resultados para a avaliação de parâmetros para o método de Lucchese (2000). As configurações utilizadas foram as mesmas para a escolha do parâmetro para o método de Marcel (1997). Diferentemente do ocorrido no método de Marcel (1997) o fator de interpolação não afeta a estimativa de rotação (Figura 3.8b). Portanto, a escolha se baseia na análise da média da REQM de translação ($med(\Delta x, \Delta y)$) e da média do TE. A Figura 3.8d destaca a relação entre esses dois critérios utilizando o conceito de dominância. Os três fatores que apresentam melhor relação entre os dois critérios possuem valores de, 3, 4 e 5. Pelo fato do valor 5 apresentar menor média de erro de translação, esse é o valor escolhido para o fator de interpolação das TF da imagens de BR a serem registradas para o método de Lucchese (2000).

Tabela 3.4: Resultado da avaliação do fator de interpolação das TF das imagens de BR para o método de Lucchese (2000), apresentando: média da REQM para os movimentos horizontais de translação (Δx); média da REQM para os movimentos verticais de translação (Δy); média entre Δx e Δy ; média da REQM para os movimentos de rotação (θ), e; média do TE.

Critérios	Fator de Interpolação														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Δx	14.52	14.11	13.70	13.50	13.23	13.52	13.50	13.39	13.30	13.45	13.29	13.26	13.49	13.41	13.16
Δy	12.64	12.90	11.81	11.79	11.69	11.61	11.61	11.55	11.52	11.50	11.57	11.54	11.86	11.61	11.58
$med(\Delta x, \Delta y)$	13.58	13.50	12.75	12.65	12.46	12.57	12.55	12.47	12.41	12.48	12.43	12.40	12.67	12.51	12.37
θ	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52	69.52
Tempo	0.85	0.89	0.99	1.09	1.24	1.42	1.58	1.76	2.00	2.11	2.29	2.46	2.01	2.07	2.16

Para o método de Vandewalle (2006) dois parâmetros são analisados. O primeiro é o tamanho do raio que será utilizado nas imagens de BR para estimar a rotação entre elas. Como esse parâmetro está relacionado apenas com o registro da rotação a Tabela 3.5 e a Figura 3.9 mostram os resultados apenas para a média da REQM de registro de rotação e a média do TE.

O erro de registro de rotação possui uma queda brusca no raio com valor igual a 0.15 e depois cresce suavemente. Já o TE tem uma característica com maior variação. Relacionando dois critérios utilizando os conceitos de dominância (Figura 3.9c) é possível notar que a melhor relação entre os critérios é o valor 0.15. Portanto, esse valor foi escolhido como parâmetro de raio para o método de Vandewalle (2006).

O segundo parâmetro tem o seu foco apenas na estimativa de translação, através da quantidade de componentes de baixa frequência utilizados no registro. O processo para estimar a translação é dependente da estimativa de rotação. Antes de registro de translação as imagens precisam ser rotacionadas. Portanto, o valor 0,15 escolhido para o raio é utilizado para estimar

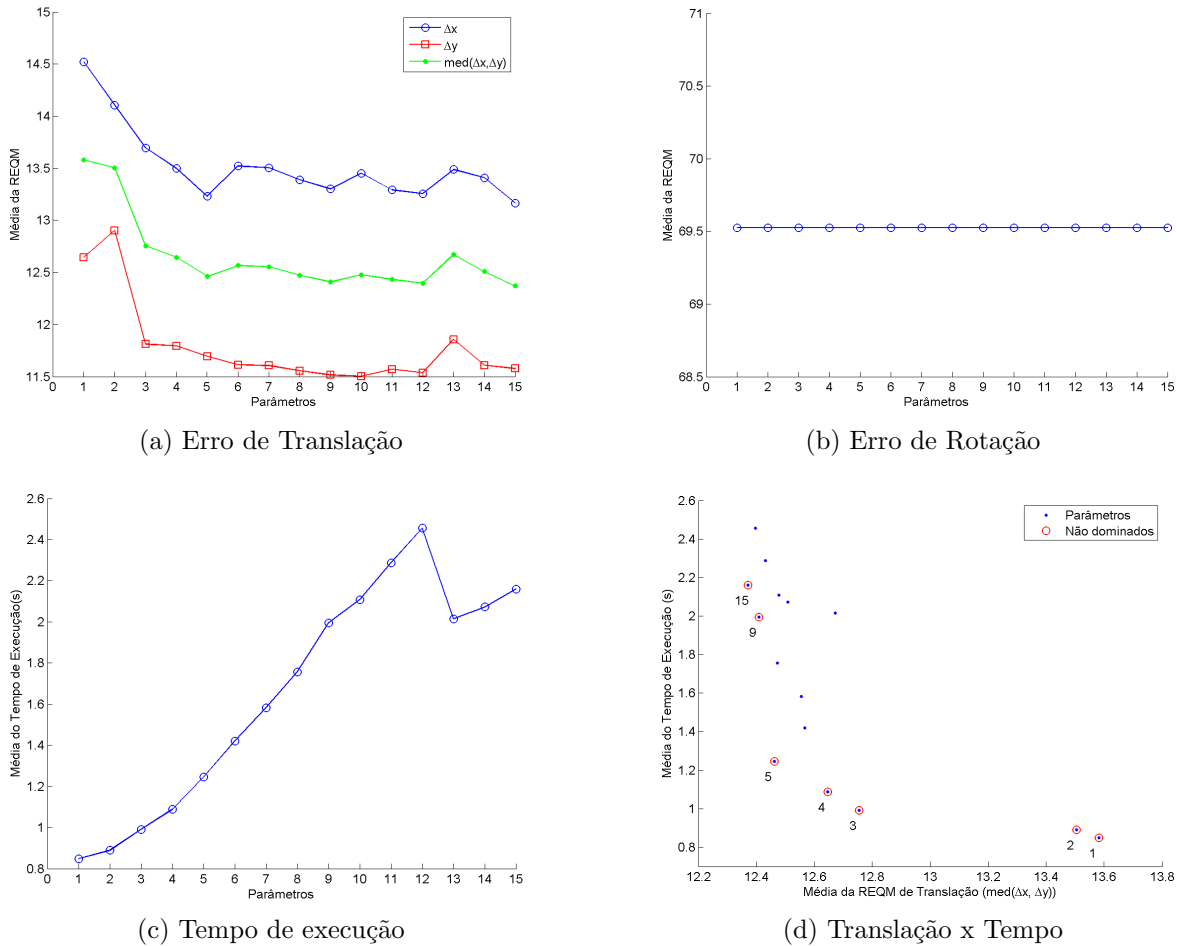


Figura 3.8: Resultados da avaliação dos fatores de interpolação aplicados ao método de Lucchese (2000). a) Média da REQM para o registro de translação; b) Média da REQM para o registro de rotação; Média do TE em segundo; d) Relação entre os erros de registro para a translação e rotação.

Tabela 3.5: Resultado da avaliação dos valores de raio para estimar a rotação para o método de Vandewalle (2006), apresentando: média da REQM para os movimentos de rotação (θ), e; média do TE.

Critérios	Raio																		
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
θ	27.26	27.26	1.96	2.21	2.52	2.95	3.29	3.85	4.32	4.51	4.54	4.45	4.37	4.22	4.18	4.14	4.12	4.12	3.94
Tempo	1.12	1.09	1.64	3.03	2.51	5.53	2.41	6.27	6.36	6.45	6.52	6.52	6.55	6.55	6.57	6.60	5.43	4.00	2.45

a rotação antes do registro de translação. A Tabela 3.6 e a Figura 3.10 mostram os resultados apenas para a média da REQM de registro de translação e a média do TE. Ao aumentar a quantidade de componentes de baixa frequência o erro de translação tende a diminuir, chegando ao mínimo em 13. Já o TE apresenta maior variação entre os parâmetros. Relacionando os resultados da média da REQM de translação ($med(\Delta x, \Delta y)$) e da média do TE (Figura 3.10c),

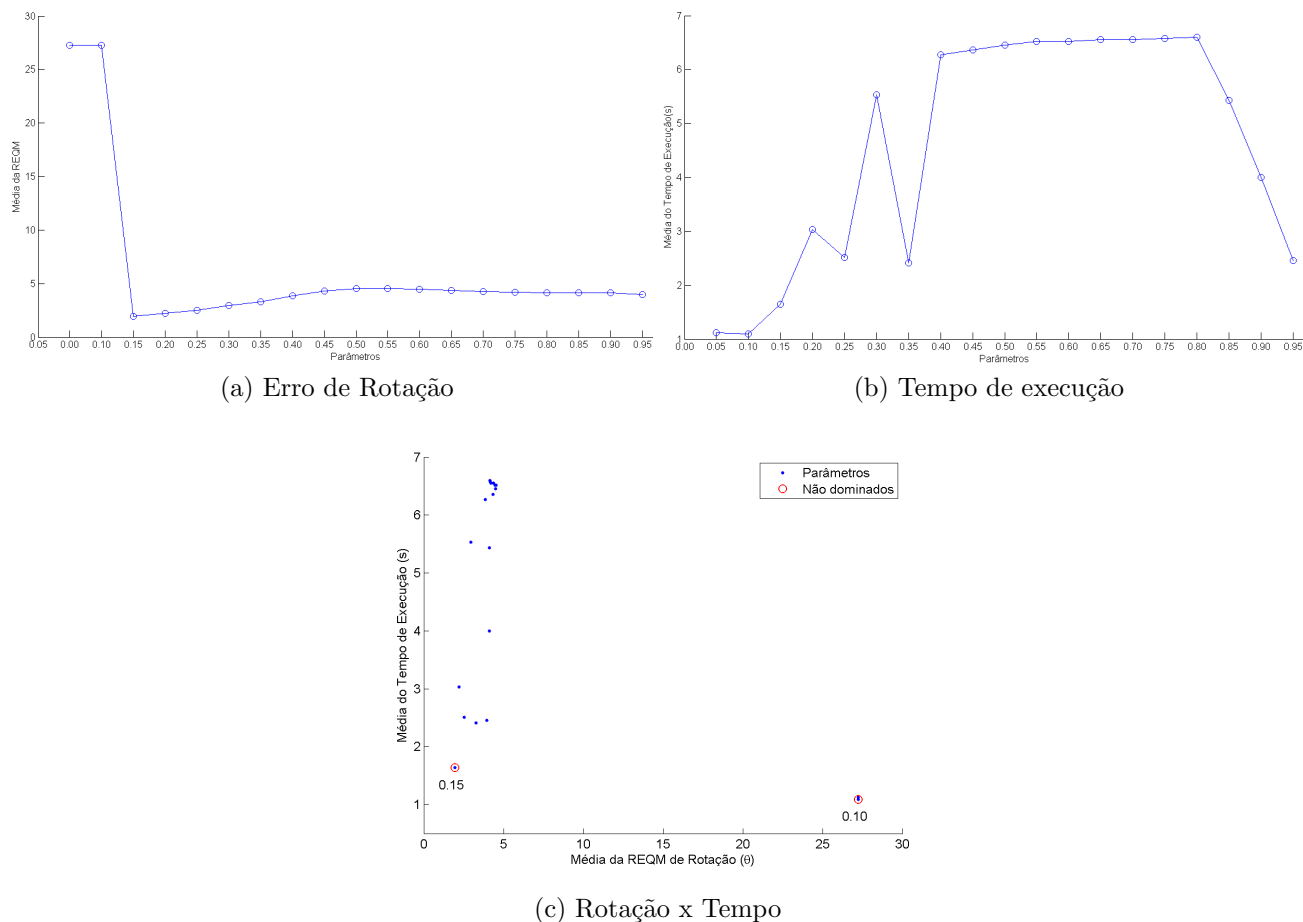


Figura 3.9: Resultados da avaliação dos valores de raio aplicados ao registro de rotação para o método de Vandewalle (2006). a) Média da REQM para no registro de rotação; b) Média do TE em segundos; c) Relação entre os erros de registro de rotação e TE

os valores não dominados, que possuem a melhor relação entre os critérios são, 6, 8 e 10. Uma vez que o valor 6 tem a menor média de TE e apresenta média de erro inferior a 0.5, esse foi o valor escolhido como a quantidade de componentes de baixa frequência nos testes para o método de Vandewalle (2006).

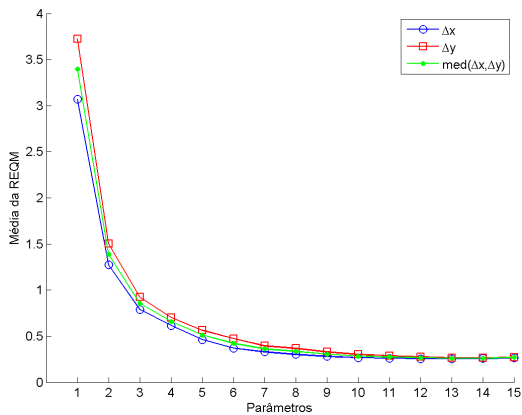
Portanto para os métodos de Marcel (1997) e Lucchese (2000) os valores do fator de interpolação da TF da imagens de BR foram definidos como 5. No método de Vandewalle (2006) o tamanho do raio para estimar a rotação entre as imagens fica definido a $0.15r$, ou seja, 15% do tamanho do raio da imagem. E para estimar a translação das imagens de BR, a quantidade de componentes de baixa frequência utilizados será de 6.

Resultados Métodos de Registro

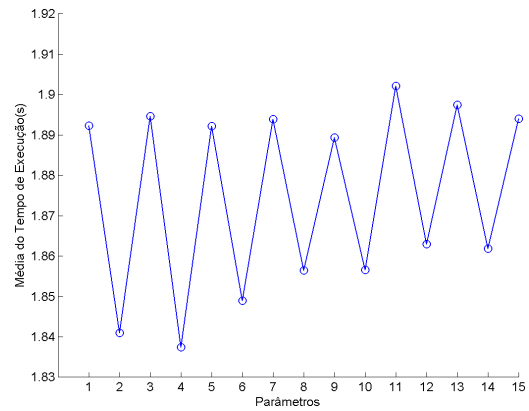
Cada um dos métodos avaliados foram testados para as 128 imagens submetidas aos 6 casos de teste propostos. As métricas utilizadas para avaliar os métodos de registro foram a REQM e o tempo de execução. A REQM foi utilizada para medir o erro entre, os valores registrados

Tabela 3.6: Resultado da avaliação da quantidade de componentes de baixa frequência para estimar a translação para o método de Vandewalle (2006), apresentando: média da REQM para os movimentos horizontais de translação (Δx); média da REQM para os movimentos verticais de translação (Δy); média entre Δx e Δy , e; média do TE.

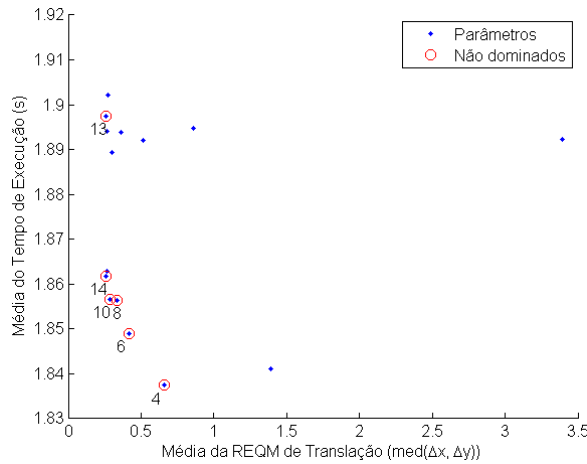
Critérios	Quantidade de componentes de baixa frequência														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Δx	3.07	1.27	0.79	0.62	0.46	0.37	0.33	0.30	0.28	0.27	0.26	0.26	0.25	0.26	0.26
Δy	3.72	1.50	0.92	0.70	0.56	0.47	0.39	0.37	0.33	0.30	0.29	0.27	0.26	0.27	0.27
$med(\Delta x, \Delta y)$	3.40	1.39	0.86	0.66	0.51	0.42	0.36	0.33	0.30	0.28	0.27	0.26	0.26	0.26	0.27
Tempo	1.89	1.84	1.89	1.84	1.89	1.85	1.89	1.86	1.89	1.86	1.90	1.86	1.90	1.86	1.89



(a) Erro de Translação



(b) Tempo de execução



(c) Translação x Tempo

Figura 3.10: Resultados da avaliação dos valores de raio aplicados ao registro de rotação para o método de Vandewalle (2006). a) Média da REQM para no registro de translação; b) Média do TE em segundos; c) Relação entre os erros de registro de translação e TE

pelos métodos nas estimativas de translação e rotação, e os valores utilizados nos operadores de transformação aplicados nas imagens de referência de AR para gerar as imagens de BR. Os

métodos que apresentam menor valor de REQM são os melhores, pelo fato de ter uma maior aproximação dos valores utilizados nos operadores de transformação ao construir as imagens de BR. Os resultados obtidos nesta etapa estão descritos a seguir. Os resultados foram agrupados pelos casos de teste. Os números apresentados representam a média dos resultados obtidos, para as 128 imagens, no caso de teste especificado.

Ao avaliar os resultados obtidos no registro de rotação, é possível notar na Tabela 3.7 que o método de Keren (1988) é superior para todos os casos de teste. O método de Marcel (1997) possui resultados semelhantes ao Vandewalle (2006) nos três primeiros casos de teste mas, a partir do caso de teste 4 os resultados são inferiores. Portanto, o método de Vandewalle (2006) é considerando o segundo melhor método neste critério.

A Figura 3.11 compara os resultados dos métodos de Keren (1988) e Vandewalle (2006). Os gráficos estão separados por caso de teste, e mostram a REQM obtida em cada uma das 128 imagens usadas nos testes. É possível observar que o método de Keren (1988) é realmente superior.

Dentre os casos de teste avaliados, o caso 5 é o que apresenta maiores valores de erro (Tabela 3.7). É possível observar que na figura Figura 3.11e que o método de Keren (1988) apresenta um resultado atípico, fora de padrão, para a imagem identificada com o número 70 (Figura 3.12). O valor da REQM, nesta imagem, para o método de Keren (1988) foi = 14.61. Foi a única imagem que o método de Vandewalle (2006) que apresentou resultado superior ao Keren (1988). Esse valor atípico é o causador do alto valor de desvio padrão para o método Keren (1988). O mesmo comportamento também pode ser percebido no caso de teste 6 (Figura 3.11f) porém em menores proporções, não tendo grande influência no desvio padrão como no caso 5.

Tabela 3.7: Comparação da média da REQM e o seu desvio padrão na estimação dos movimentos de rotação (θ), pelos métodos de registro avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.

Caso	FE	Keren	Marcel	Lucchese	Vandewalle
1	2	0.20(0.13)	0.71(0.00)	58.97(11.33)	0.71(0.03)
2	2	0.22(0.19)	0.65(0.38)	46.72(14.02)	0.65(0.38)
3	4	0.65(0.33)	1.87(0.03)	68.06(6.84)	1.87(0.02)
4	4	0.76(0.39)	3.53(15.56)	64.95(7.55)	2.16(0.37)
5	6	1.29(1.30)	3.83(8.90)	69.36(5.62)	3.03(0.01)
6	6	0.52(0.31)	1.77(2.99)	51.43(5.35)	1.41(0.09)

Na Tabela 3.8 são descritos os resultados obtidos para a estimativa dos movimentos de translação para cada um dos métodos avaliados. O método de Keren (1988) obteve os melhores resultados para os caso teste 1, 2 e 4, e o método de Vandewalle (2006) se destacou no caso 5. Nos casos 3 e 6 não se pode afirmar que o método de Keren (1988) é superior ao Vandewalle (2006) devido a alta variabilidade dos resultados observada no valor do desvio padrão. Os outros dois métodos se mostram bem inferiores.

A Figura 3.13 e Figura 3.14 mostram, respectivamente, os resultados individuais dos métodos de Keren (1988) e Vandewalle (2006) para estimar a translação no eixo x e no eixo y . Nos casos de teste 3 e 6, tanto para os deslocamentos no eixo x quanto no eixo y , em semelhança ao ocorrido no caso 5 do registro de rotação, o método de Keren (1988) apresentou um resultado fora de padrão, elevando o valor da média REQM e desvio padrão. No caso de teste 5,

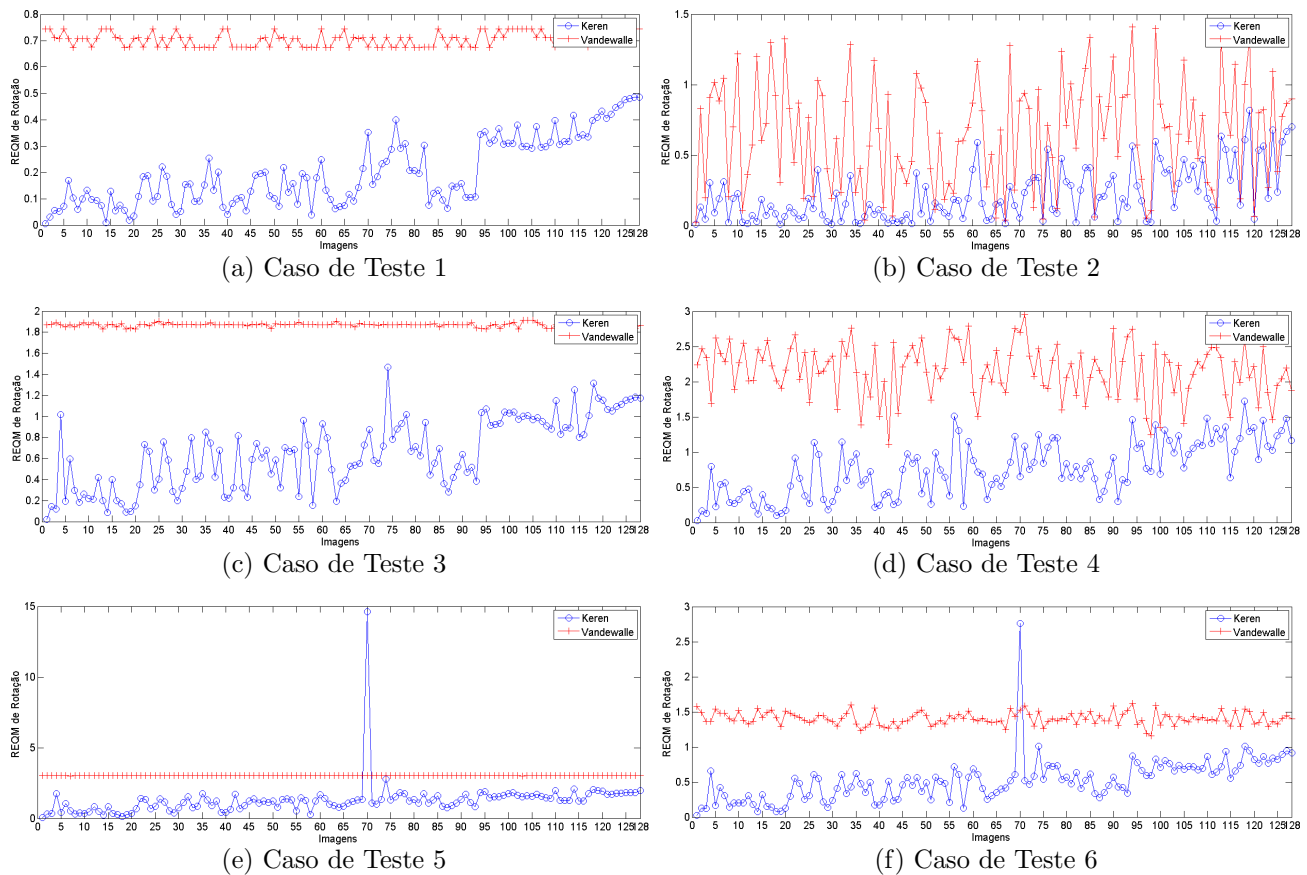


Figura 3.11: Resultados da REQM no registro de rotação dos métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste.

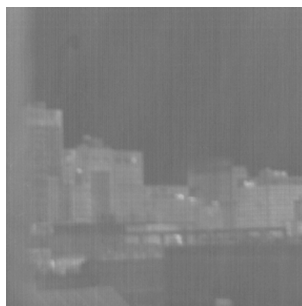


Figura 3.12: Imagem de número 70 utilizada nos testes. Imagem que originou o conjunto de imagens de BR, que ao ter a rotação estimada pelo método de Keren (1988), no caso de teste 5, apresentou um resultado atípico.

tanto na Figura 3.13e quanto na Figura 3.14e, é possível notar que pelo menos quatro valores se mostraram atípicos ao padrão formado no método de Keren (1988) (Figura 3.15 e Figura 3.16). Esses valores contribuem significativamente para o resultado da média e desvio padrão do método de Keren (1988), tornando o resultado neste caso de teste, inferior ao método de

Vandewalle (2006).

Observando os resultados obtidos em cada um dos casos de teste, é possível notar que o método de Keren (1988), no geral, é superior ao método de Vandewalle (2006) para os registros de translação. Essa conclusão é baseada na observação que o método de Keren (1988) possui melhores médias de resultados na maioria dos testes. Apenas em alguns poucos casos o método apresenta alguns resultados atípicos.

Tabela 3.8: Comparação da média da REQM e o seu desvio padrão na estimação dos movimentos de translação (Δx , Δy), pelos métodos de registro avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.

Caso	FE	Keren		Marcel		Lucchese		Vandewalle	
		Δx	Δy	Δx	Δy	Δx	Δy	Δx	Δy
1	2	0.14(0.18)	0.16(0.16)	3.56(10.01)	11.07(22.84)	25.48(22.18)	25.56(22.89)	0.31(0.27)	0.42(0.26)
2	2	0.25(0.22)	0.24(0.20)	3.44(11.47)	7.81(21.42)	20.18(25.16)	14.79(20.09)	0.35(0.34)	0.45(0.42)
3	4	0.31(1.10)	0.47(1.87)	13.09(10.70)	12.71(10.30)	17.38(10.63)	16.56(10.72)	0.36(0.26)	0.48(0.28)
4	4	0.28(0.21)	0.35(0.30)	12.44(11.40)	12.53(11.56)	17.49(11.09)	17.33(11.18)	0.43(0.28)	0.52(0.31)
5	6	0.42(0.84)	0.79(2.53)	11.43(7.83)	11.02(7.05)	12.74(7.12)	12.44(6.97)	0.37(0.21)	0.46(0.23)
6	6	0.23(0.36)	0.34(0.76)	7.33(6.23)	9.19(8.87)	15.57(7.49)	14.48(6.44)	0.30(0.19)	0.39(0.22)

Levando em consideração a média do TE, o método de Keren (1988) foi superior na maioria dos casos (Tabela 3.9). O método de Vandewalle (2006) foi, superior no caso de teste 5. No caso de teste 3, devido a variabilidade apresentada pelo método de Keren (1988) não é possível afirmar que este é superior ao Vandewalle (2006). Os resultados de cada uma das imagens em cada um dos testes, para os métodos de Keren (1988) e Vandewalle (2006), também estão demonstrados na Figura 3.17.

Tabela 3.9: Comparação da média e o seu desvio padrão, do TE gasto para estimar o registro das imagens, pelos métodos de registro avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.

Caso	FE	Keren	Marcel	Lucchese	Vandewalle
1	2	0.18(0.03)	0.69(0.02)	0.50(0.05)	0.33(0.04)
2	2	0.06(0.01)	0.24(0.01)	0.18(0.02)	0.17(0.02)
3	4	0.37(0.07)	0.86(0.03)	0.84(0.10)	0.38(0.04)
4	4	0.17(0.03)	0.40(0.01)	0.40(0.05)	0.20(0.03)
5	6	0.62(0.12)	0.76(0.04)	1.25(0.16)	0.56(0.06)
6	6	0.23(0.04)	0.49(0.02)	0.53(0.06)	0.27(0.03)

Com base nos resultados apresentados o método de Keren (1988) foi escolhido como o método indicado para o registro das imagens. O método obteve destaque nos três critérios avaliados. Para o registro de rotação, o método obteve melhores resultados de média da REQM em todos os casos de teste. No caso de teste 5 o método apresentou um alto valor de desvio padrão. Porém, foi demonstrado que esse valor foi causado por resultado atípico em todos as imagens avaliadas. No registro de translação, apenas para o caso de teste 5 o método de Keren (1988) não foi superior quanto a média da REQM. No entanto, esse resultado, em semelhança do ocorrido no registro de rotação para o mesmo caso de teste, é justificado pelos resultados atípicos ocorridos em menos de 4% das imagens avaliadas. Quanto ao TE, foi superior em 4 dos 6 casos de teste avaliados.

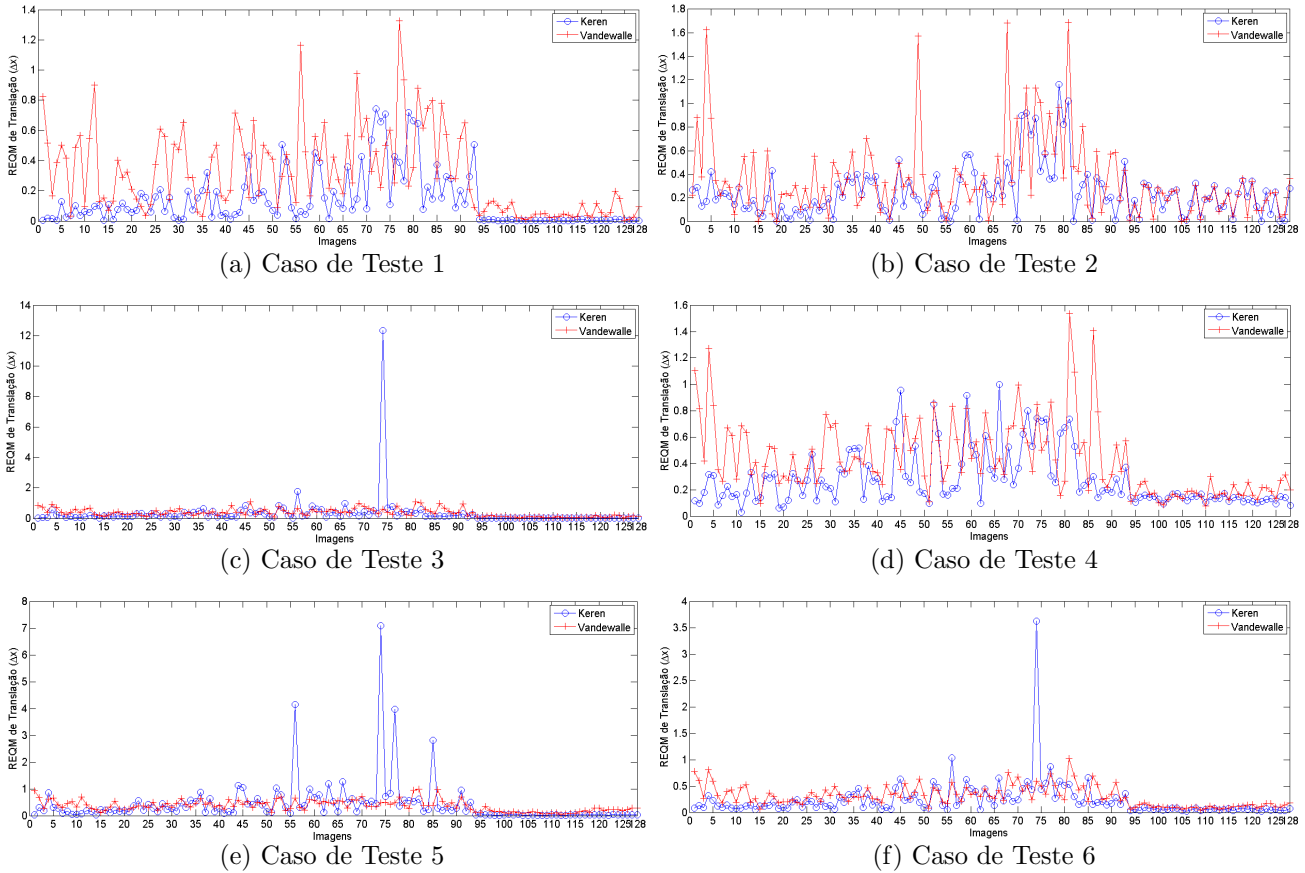


Figura 3.13: Resultados da REQM no registro de translação no eixo x dos métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste.

Resultados Métodos de Reconstrução

Como a fase de reconstrução é consecutiva à fase de registro, os métodos avaliados tem como entrada os valores dos deslocamentos entre as imagens do conjunto. Os valores utilizados para gerar cada uma das imagens de BR foram passados aos métodos de reconstrução. O objetivo é que os resultados não sejam afetados por erros oriundos do registro das imagens, prejudicando a análise dos métodos.

Assim como na avaliação dos métodos de registro, os resultados foram agrupados por caso de teste. As tabelas a seguir (Tabela 3.11, Tabela 3.10, Tabela 3.12) mostram os resultados obtidos para as seguintes métricas de avaliação: RSRP, CCP e TE.

Os resultados obtidos em relação a RSRP (Tabela 3.10) mostram que o método G-PMSR é superior na maioria dos casos de teste. O único caso em que o resultado foi diferente foi para o caso de teste 2, onde o IBP e o RSR foram superiores. Os experimentos com o CCP (Tabela 3.11) tiveram resultados semelhantes aos da RSRP. É possível observar que o G-PMSR teve, no geral, a melhor performance. O único caso onde o G-PMSR também não se destacou foi no caso de teste 2, onde os melhores resultados foram obtidos pelos métodos IBP e RSR.

Através dos resultados é possível perceber a influência da relação FE e quantidade de ima-

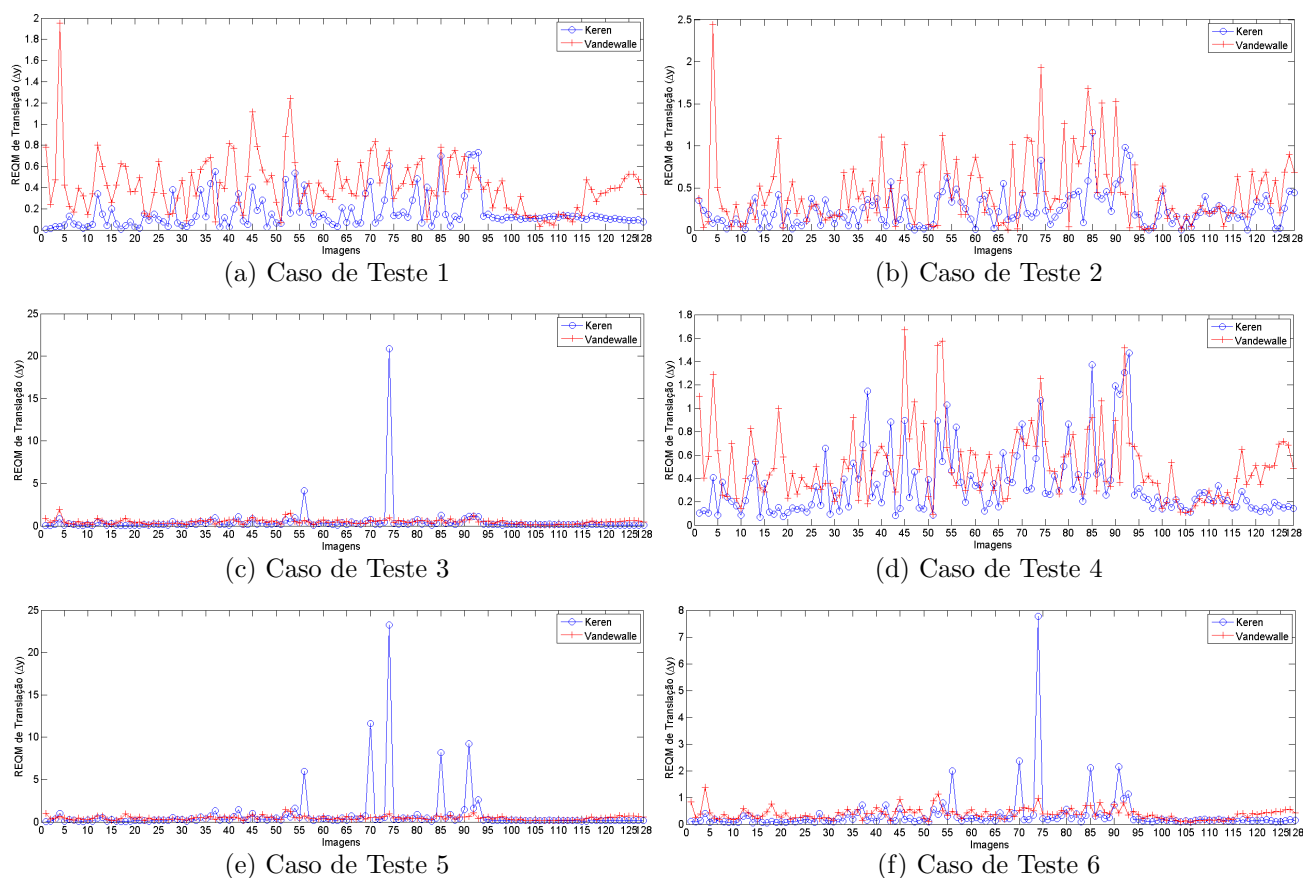


Figura 3.14: Resultados da REQM no registro de translação no eixo y dos métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste.

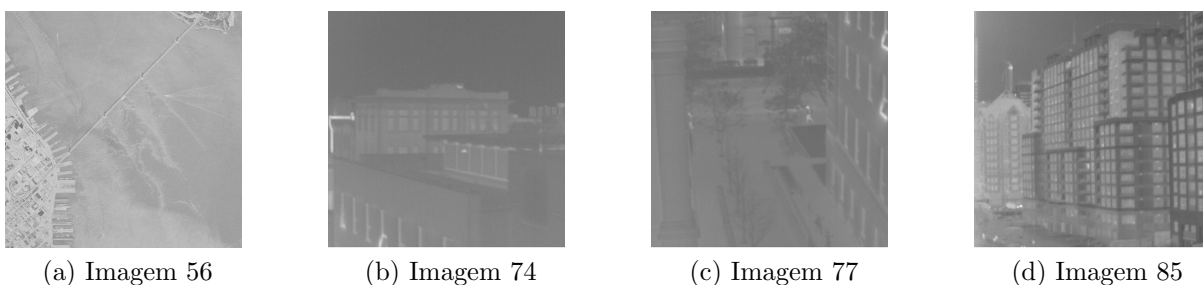


Figura 3.15: Imagens que originaram conjuntos de imagens de BR, que ao terem as translações no eixo x estimadas pelo método de Keren (1988), no caso de teste 5, apresentaram resultado atípico.

gens de BR disponíveis no conjunto de teste. Nos casos de teste com FE baixo, os resultados apresentados foram melhores do que nos casos com FE elevado. Observando a Tabela 3.11, ao comparar os resultados de cada método, individualmente, entre os casos de teste 1 e 6, é possível notar que, a queda de performance para o método G-PMSR foi de aproximadamente

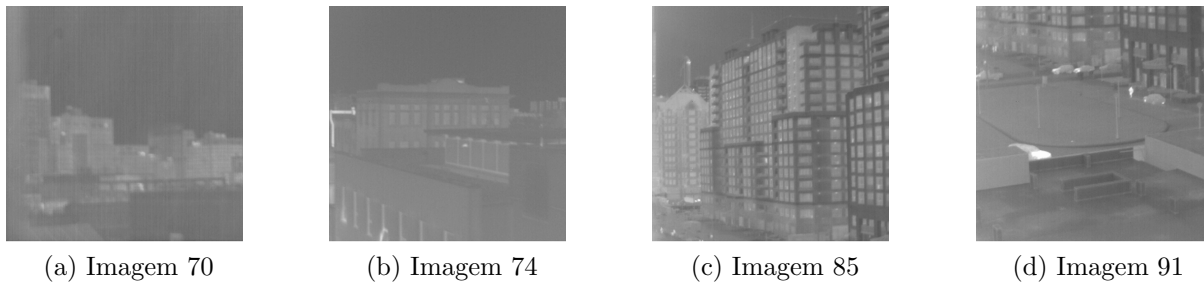


Figura 3.16: Imagens que originaram conjuntos de imagens de BR, que ao terem as translações no eixo y estimadas pelo método de Keren (1988), no caso de teste 5, apresentaram resultado atípico.

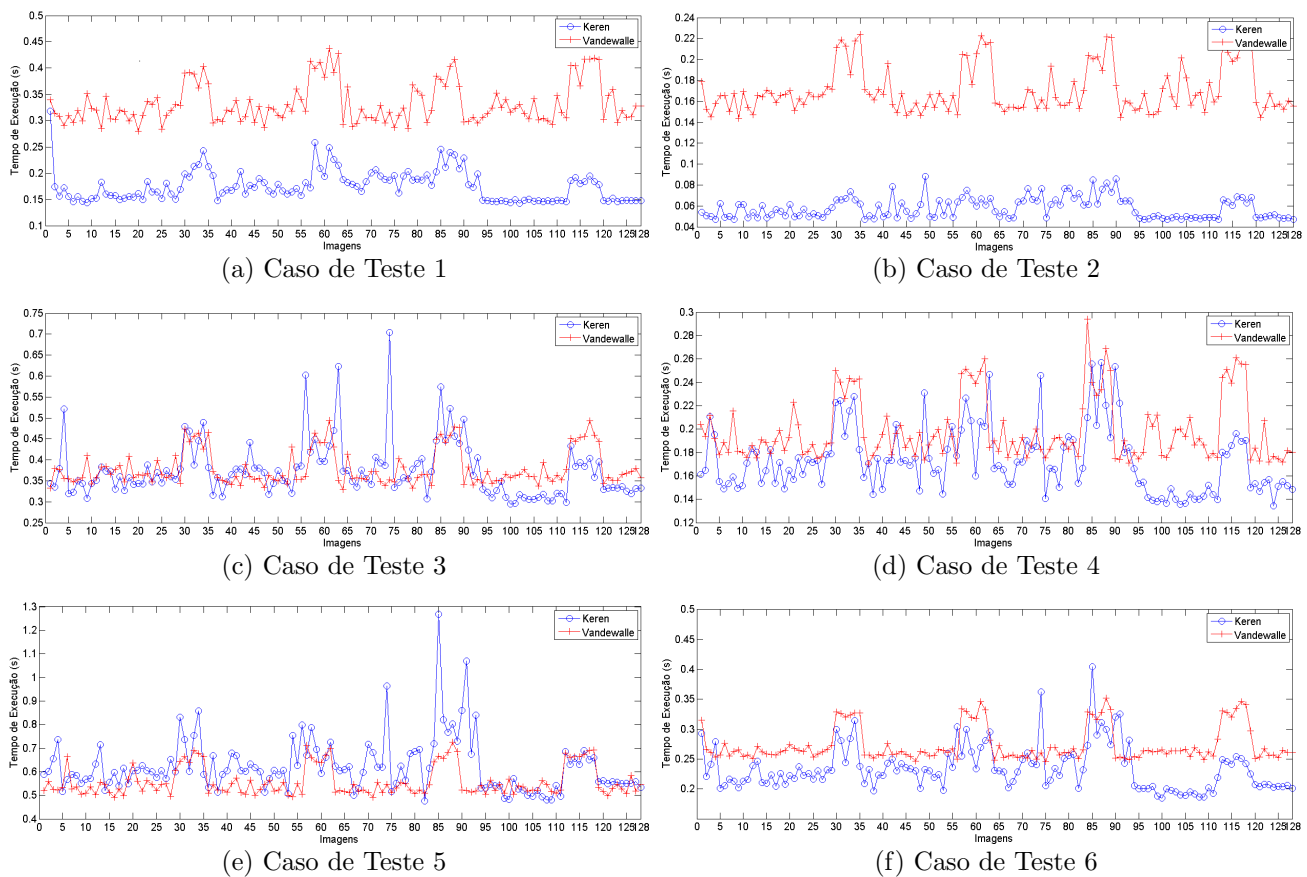


Figura 3.17: Resultados do TE para os métodos de Keren (1988) e Vandewalle (2006) para as 128 imagens separadas por caso de teste.

de 14%, já nos outros métodos, a queda foi maior do que 20%. Comparando dois a dois os casos com um mesmo valor de FE, na Tabela 3.11, em todos os casos, os resultados que possuem menor quantidade de imagens obtiveram piores resultados. No geral, há uma depreciação em torno de 5% nos resultados.

Tabela 3.10: Comparação da média da RSRP e o seu desvio padrão obtido na estimação da imagem de AR pelos métodos de reconstrução avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.

Caso	FE	Qtd. Img.	IBP	TV	RSR	NC	G-PMSR
1	2	4	30.03(4.73)	32.12(4.59)	30.04(4.75)	30.68(4.30)	32.98(5.00)
2	2	2	29.88(4.72)	27.69(3.46)	29.88(4.72)	28.99(3.79)	28.15(3.62)
3	4	16	25.62(3.16)	26.43(3.35)	25.69(3.43)	25.58(3.49)	28.05(3.49)
4	4	8	24.99(3.47)	24.80(3.15)	24.95(3.52)	24.52(3.56)	25.98(3.18)
5	6	36	21.57(4.05)	23.67(4.17)	21.93(5.56)	23.23(3.41)	26.10(3.34)
6	6	18	22.28(2.63)	22.00(5.05)	23.14(3.50)	22.59(3.42)	24.54(3.31)

Tabela 3.11: Comparação da média do CCP e o seu desvio padrão obtido na estimação da imagem de AR pelos métodos de reconstrução avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.

Caso	FE	Qtd. Img.	IBP	TV	RSR	NC	G-PMSR
1	2	4	0.94(0.06)	0.96(0.04)	0.94(0.06)	0.94(0.06)	0.96(0.04)
2	2	2	0.94(0.06)	0.90(0.09)	0.94(0.06)	0.91(0.08)	0.90(0.09)
3	4	16	0.86(0.09)	0.88(0.08)	0.87(0.09)	0.82(0.16)	0.91(0.08)
4	4	8	0.84(0.10)	0.84(0.11)	0.84(0.10)	0.78(0.19)	0.86(0.10)
5	6	36	0.72(0.17)	0.80(0.12)	0.74(0.16)	0.74(0.22)	0.87(0.10)
6	6	18	0.74(0.16)	0.74(0.16)	0.78(0.12)	0.72(0.22)	0.82(0.12)

Com relação ao TE (Tabela 3.12) é possível notar que a tarefa de reconstrução é, em si, custosa. O método que obteve os melhores resultados nesse critério foi, indiscutivelmente, o IBP proposto por Irani (1991). Em todos os casos o IBP apresentou TE bastante inferior comparando-o com os outros métodos. É relevante notar que o método G-PMSR, que obteve em geral o menor erro, é também, aquele que possui o maior custo computacional para a resolução do problema de reconstrução da imagem de AR.

Tabela 3.12: Comparação da média do TE, em segundos, e o seu desvio padrão obtido na estimação da imagem de AR pelos métodos de reconstrução avaliados, levando em consideração 6 casos de teste, executados para as 128 imagens.

Caso	FE	Qtd. Img.	IBP	TV	RSR	NC	G-PMSR
1	2	4	1.11(0.59)	71.11(32.56)	15.20(6.16)	62.09(24.76)	76.52(36.46)
2	2	2	0.95(0.57)	67.17(35.47)	14.95(6.00)	33.44(12.29)	79.12(37.07)
3	4	16	12.93(5.74)	71.23(31.08)	70.45(39.92)	58.29(21.89)	112.69(63.41)
4	4	8	3.40(2.91)	70.02(36.25)	30.03(31.71)	31.94(11.48)	84.96(39.38)
5	6	36	28.96(26.72)	106.40(113.46)	102.65(106.96)	59.70(23.81)	137.64(64.47)
6	6	18	20.07(8.06)	110.29(130.09)	80.70(60.76)	34.26(12.52)	85.99(52.58)

A título de exemplo as Figuras 3.18 e 3.19 esboçam os resultados dos métodos de reconstrução para imagens de teste. Os casos de teste com FE maiores geram, também, maiores dificuldades para os métodos de reconstrução, por isso os métodos tendem a produzir imagens com maiores índices de degradação. A Figura 3.19 apresenta um corte de 100x100 píxeis das imagens da Figura 3.18. É possível notar que o método G-PMSR produz resultados mais homogêneos, com uma melhor capacidade de recuperar as regiões de alta frequência na imagem, se comparado com os outros métodos avaliados. Os três primeiros métodos, IBP, TV e RSR quando aplicados a um FE maior produzem imagens com maior efeito de *aliasing*, as imagens

ficam com aparência pixelizada. Já os métodos NC e G-PMSR produzem imagens mais suaves. Porém, nota-se a presença de efeito de ruído nos resultados da NC.

Levando em consideração todas as métricas avaliadas, o método de reconstrução sugerido é o proposto por Maiseli (2014), o G-PMSR. O G-PMSR teve os melhores resultados para 2 das 3 métricas avaliadas. A única métrica em que o método se mostrou deficiente, em relação aos outros avaliados, foi o TE. Em problemas cujo custo computacional é o fator principal a ser observado, a opção indicada analisando o resultado dos experimentos seria o método IBP. Porém, a escolha pelo IBP não é recomendada para valores elevados de FE, pois a utilização do método nessas condições torna-se inviável, uma vez que não é capaz de produzir resultados com a precisão esperada.

3.3 Experimentos com imagens reais

3.3.1 Metodologia

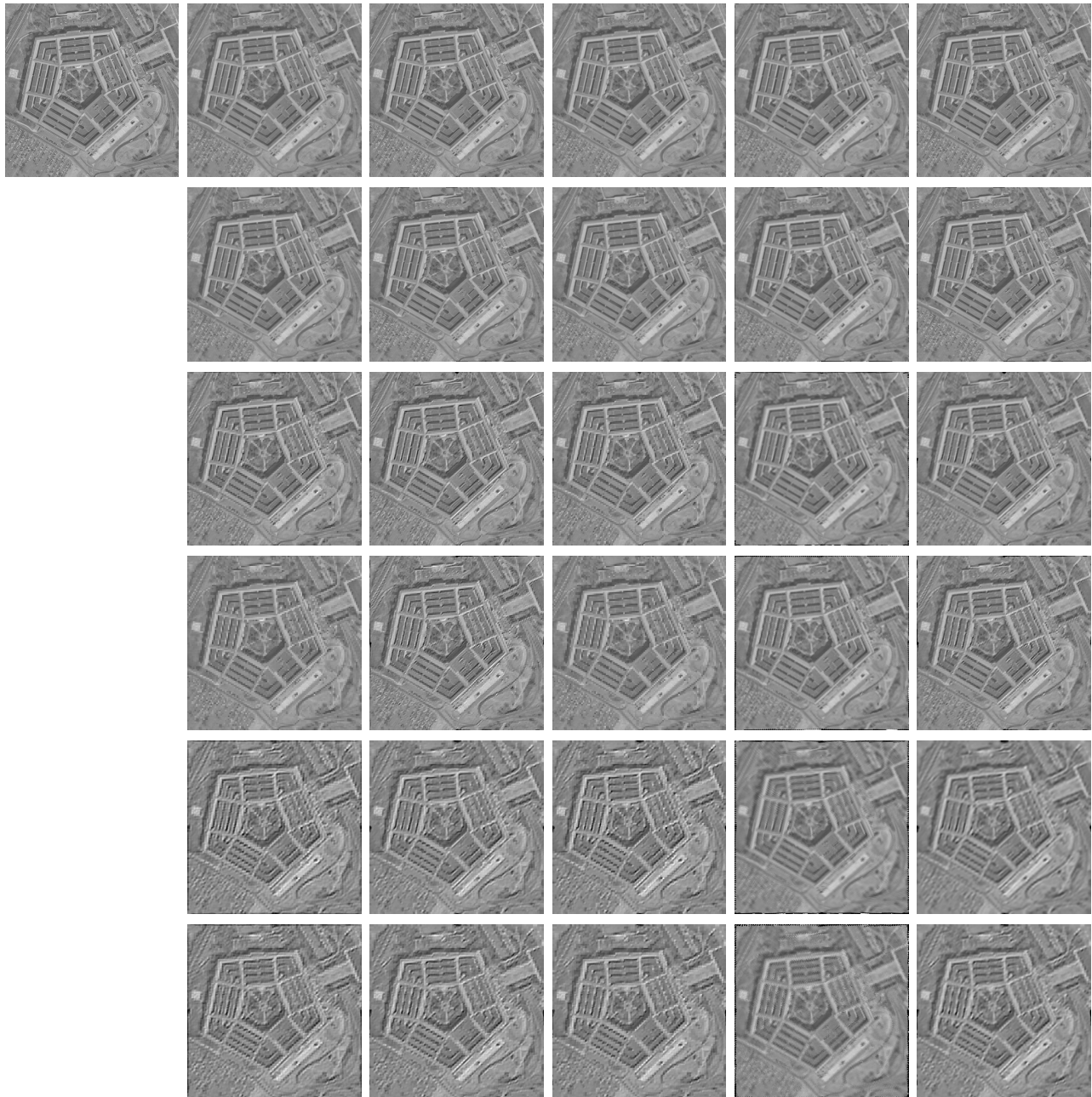
O experimento com imagens simuladas teve como objetivo definir quais são os melhores métodos nas etapas de registro e reconstrução. Para o registro, o método que obteve os melhores resultados foi o método proposto por Keren (1988). Na etapa de reconstrução o método em superioridade foi o proposto por Maiseli (2014). Assim, esses dois métodos podem ser combinados para sua aplicação em problemas de SR.

Uma vez escolhida a melhor combinação de técnicas, o principal objetivo dos experimentos com imagens reais é validar os resultados obtidos nos experimentos com as imagens simuladas. As duas técnicas serão utilizadas em conjunto para gerar uma imagem de AR, a partir de imagens de BR. Diferentemente dos experimentos anteriores, no caso dos experimentos com imagens reais, não se sabe o processo ao qual as imagens de BR foram formadas. Aqui não se tem informações sobre quais as transformações envolvidas no processo de formação das imagens, e nem as degradações sofridas pelas imagens de baixa resolução ao amostrar a cena no domínio contínuo. Sete diferentes conjuntos de imagens foram selecionados para essa etapa de experimentos.

Cada conjunto é inicialmente aplicado ao método de Keren (1988) para realizar o registro das imagens. As informações de registro são então utilizadas como entrada para o método de Maiseli (2014) para produzir uma hipótese da imagem de AR com FE igual a 6. Como não existe um padrão de comparação para as imagens de AR, a avaliação será restrita a aspectos qualitativos selecionando regiões de interesse nas imagens para análise. Assim, também são mostrados os resultados para o método de Interpolação Bicubica (BI). O objetivo de usar a BI não é para compará-las, ou mesmo tentar afirmar que a técnica de SR é superior. A utilização da BI é para auxílio na avaliação qualitativa dos resultados da SR.

3.3.2 Dados

Como dito anteriormente as técnicas de SR com múltiplas imagens necessitam de um conjunto de imagens de BR com características bem específicas. As imagens devem ser diferentes, porém essas diferenças devem ser mínimas a nível de sub-píxel. Esse foi um dos aspectos observados para a escolha das imagens para os experimentos com imagens reais. Como não é



(a) Original

(b) IBP

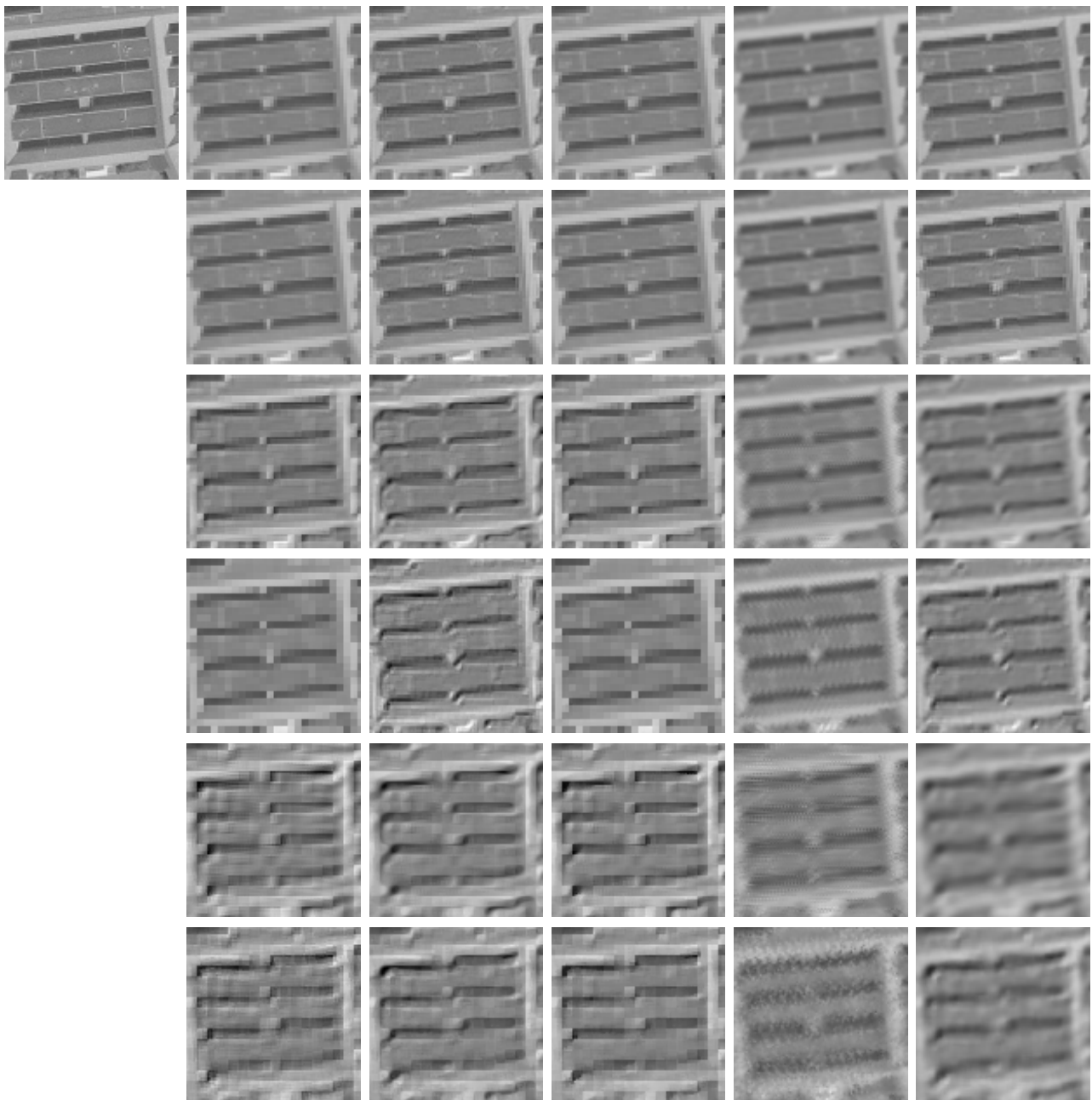
(c) TV

(d) RSR

(e) NC

(f) G-PMSR

Figura 3.18: Resultados do métodos de reconstrução de uma imagem de teste, para todos os 6 casos de teste (Tabela 3.1), representados em cada linha, sendo a primeira linha o caso de teste 1 e a última o caso de teste 6. Linha 1: $FE = 2$, qtd. imgs. = 4, res = 240x240; Linha 2: $FE = 2$, qtd. imgs. = 2, res = 240x240; Linha 3: $FE = 4$, qtd. imgs. = 16, res = 120x120; Linha 4: $FE = 4$, qtd. imgs. = 8, res = 120; Linha 5: $FE = 6$, qtd. imgs. = 36, res = 80x80; Linha 6: $FE = 6$, qtd. imgs. = 18, res. = 80x80;



(a) Original (b) IBP (c) TV (d) RSR (e) NC (f) G-PMSR

Figura 3.19: Recorte de uma região de interesse de 100x100 píxeis das imagens da Figura 3.18. Todos os 6 casos de teste (Tabela 3.1) estão representados em cada linha, sendo a primeira linha o caso de teste 1 e a última o caso de teste 6. Linha 1: FE = 2, qtd. imgs. = 4, res = 240x240; Linha 2: FE = 2, qtd. imgs. = 2, res = 240x240; Linha 3: FE = 4, qtd. imgs. = 16, res = 120x120; Linha 4: FE = 4, qtd. imgs. = 8, res = 120; Linha 5: FE = 6, qtd. imgs. = 36, res = 80x80; Linha 6: FE = 6, qtd. imgs. = 18, res. = 80x80;

possível afirmar que o conjunto de imagens de BR possui deslocamentos de sub-píxel entre si, parte dos conjuntos selecionados a partir de trabalhos que estão relacionados a SR com múltiplas imagens, como no caso do conjunto A, B e C (Figura 3.20). Outros conjuntos (D, E e F) as imagens são provenientes de uma sequência de vídeo. A probabilidade de haver deslocamentos de sub-píxel entres os frames de vídeo é maior. No caso específico do conjunto G, as imagens são fotos tiradas especificamente para o presente trabalho. Ao capturar as fotos, houve a precaução da movimentação mínima tanto do fotógrafo quanto da pessoa presente na cena. Outro fator que influenciou na escolha, está relacionado ao conteúdo da cena. Procurou-se encontrar cenas que permitam uma avaliação qualitativa dos resultados, como, por exemplo, textos e formas geométricas.

Um exemplo de cada uma das imagens de BR utilizadas nos 7 conjuntos avaliados é mostrado na Figura 3.20. Foram selecionados 3 conjuntos obtidos em Milanfar (2011), 3 obtidos em Sunkavalli (2012) e um produzido para a pesquisa com imagens térmicas obtidas através de um sensor FLIR com espectro de $7.5\text{-}13\ \mu\text{m}$ e resolução de 240×240 píxeis. Os conjuntos de imagens obtidos por Sunkavalli (2012) foram gerados a partir de vídeos. Os vídeos em questão possuem frames em AR. Porém, ao selecionar apenas uma região de interesse no frame de AR é possível obter uma imagem de BR. Portanto, ao selecionar a mesma posição de corte em cada um dos frames do vídeo é possível obter um conjunto de imagens de BR. A Tabela 3.13 apresenta as características de cada grupo selecionado para esta fase de experimentos.

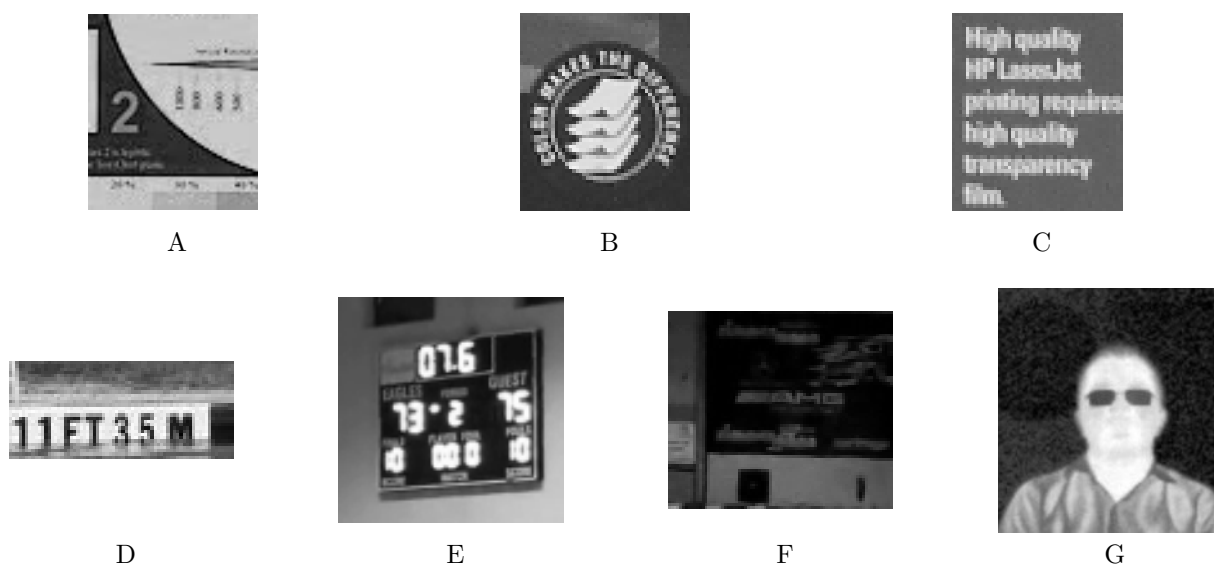


Figura 3.20: Uma imagem de BR de cada conjunto de teste a ser avaliado. A-C Milanfar (2011); D-F Sunkavalli (2012); G FLIR InfraCamSD.

3.3.3 Resultados

Os resultados obtidos para os 7 conjuntos de imagens investigados são descritos abaixo. Para cada um dos conjuntos, uma das imagens de BR do conjunto é colocada lado a lado com

Tabela 3.13: Casos de teste propostos para a primeira etapa dos experimentos.

Conjunto	Qtd. de Imagens	Resolução	Origem	Fonte
A	20	66x76	Fotos	Milanfar (2011)
B	19	49x57	Fotos	Milanfar (2011)
C	29	49x57	Fotos	Milanfar (2011)
D	13	70x30	Vídeo	Sunkavalli (2012)
E	31	90x90	Vídeo	Sunkavalli (2012)
F	26	80x70	Vídeo	Sunkavalli (2012)
G	19	175x190	Fotos	Própria

os resultados obtidos pela interpolação Bicubica e pela técnica de SR, obtida pela combinação dos métodos de Keren (1988) e Maiseli (2014). Além disso, regiões de interesse são selecionadas nos resultados para comparação.

Os resultados para o conjunto A são demonstrados na Figura 3.21. Ao observar as Figuras 3.21b e 3.21c é possível perceber uma maior nitidez na Figura 3.21c. Analisando a primeira região selecionada na imagem (Figuras 3.21d e 3.21e) é possível perceber que o resultado obtido pelo método de SR tem uma melhor resposta na recuperação das regiões de alta frequência na imagem. É possível notar que as bordas do triângulo possuem uma melhor definição na imagem obtida pela SR. Na segunda região selecionada (Figuras 3.21f e 3.21g) é possível notar, pelo resultado da SR (Figura 3.21g), o número “1000” orientado verticalmente. Observando o resultado da interpolação Bicubica não é possível discernir o número “1000” na cena.

Nos conjuntos de teste B e C é possível notar a presença de textos. Os textos presentes nas cenas servem de auxílio para avaliar os resultados obtidos. Pois já existe um conhecimento *a priori* das formas das letras. Os resultados para o conjunto B podem ser observados na Figura 3.22. Comparando os resultados da interpolação Bicubica com o método de SR é possível notar que, o texto presente na cena, torna-se mais evidente no resultado da SR (Figura 3.22c). Ao observar a Figura 3.22d e a Figura 3.22e nota-se claramente as letras contidas no recorte, auxiliando na melhor compreensão da palavra em questão “COLOR”. Na segunda região selecionada nos resultados pode-se perceber que na Figura 3.22f as letras são dificilmente percebidas. Ao passo que na Figura 3.22g é possível notar as letras “C” e “E”. O círculo que sobreescreve o texto também se mostra mais nítido no resultado da SR. É possível notar essa diferença também na Figura 3.22f e na Figura 3.22g. A cena representada pelo conjunto de teste C é um texto (Figura 3.23). O resultado da interpolação Bicubica apresenta uma imagem com efeito de *blur*. Com a SR o efeito de *ghosting* pode ser percebido. Nos cortes selecionados é possível perceber que as letras são melhor identificadas no resultado da SR. Na imagem Figura 3.23d a letra “i” é imperceptível. E no segundo corte a letra “a” é evidenciada nas Figuras 3.23f e 3.23g.

A cena do vídeo do conjunto D focaliza uma inscrição na borda de uma piscina. Assim como nos resultados apresentados pelos conjuntos B e C a resposta do método de SR para o conjunto D também produz uma imagem com maior nitidez. Observando a Figura 3.24 é possível notar que o texto da cena pode ser reconhecido. Porém, pela observação dos recortes nos resultados, a delimitação das bordas no resultado da SR é superior.

No caso do conjunto de teste E, a cena é um placar eletrônico de um jogo de basquete (Figura 3.25). Diferente dos métodos anteriores, o resultado da SR não é aparentemente superior ao

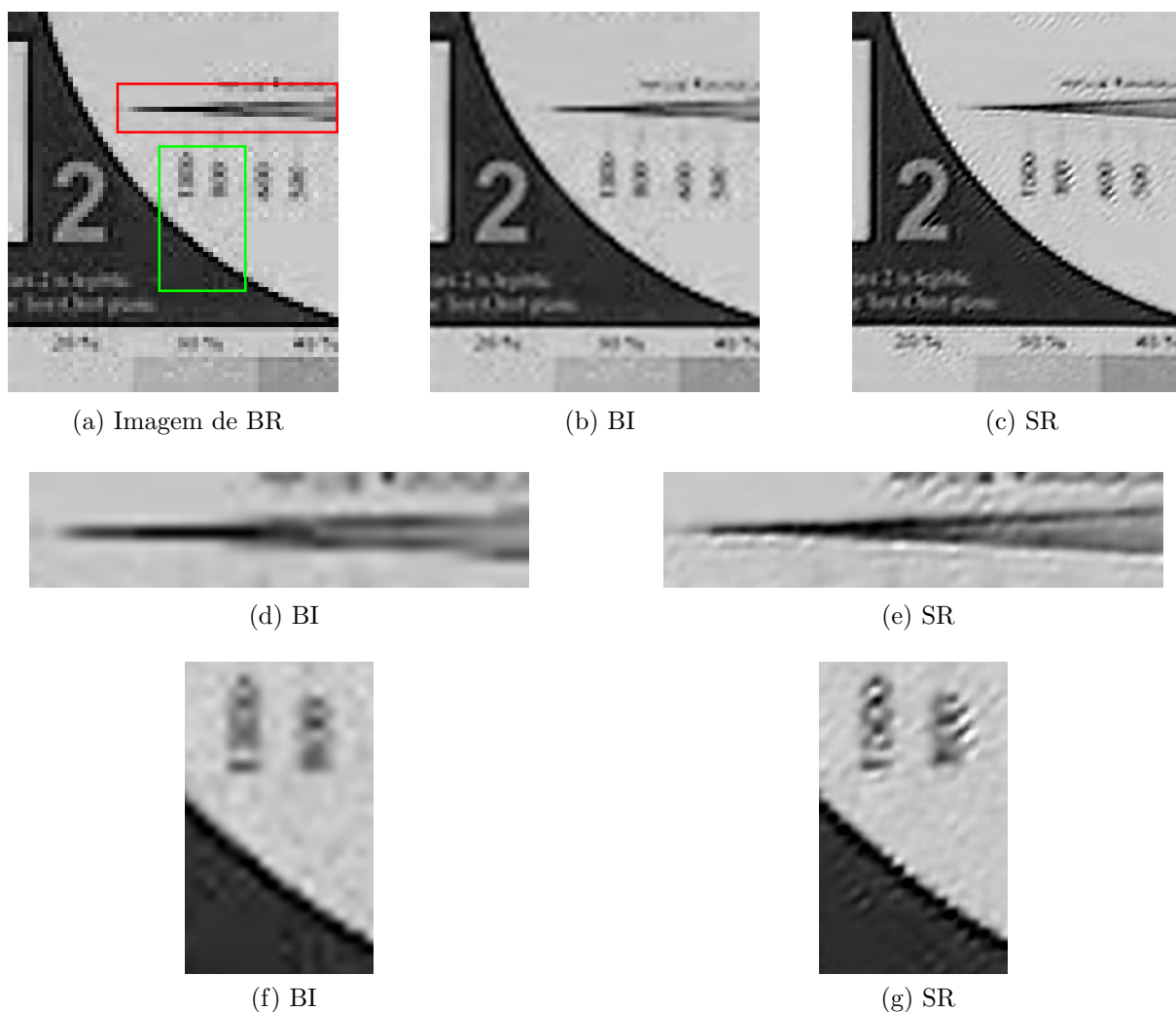


Figura 3.21: Resultado da SR para o conjunto A. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

método da interpolação Bicubica. Observando as regiões de interesse recortadas nos resultados, nota-se o mesmo padrão percebido nos experimentos anteriores. O método de SR apresenta uma melhor recuperação das regiões de alta frequência. A linha formada pela base do placar eletrônico na Figura 3.25d possui algumas distorções. Esse efeito não acontece no resultado da SR Figura 3.25e. O número “6” está suavemente melhor definido na Figura 3.25g se comparado com a Figura 3.25f. A partir destas observações pode se dizer que o resultados da SR foi um pouco melhor do que a interpolação Bicubica.

O conjunto de imagens extraídas do vídeo do caso de teste F apresenta um muro de propagandas ao fundo de um circuito de kart. Nos dois resultados é possível notar que no centro da imagem se encontra a inscrição “AMG”. Selecionando a região que contorna a inscrição, em

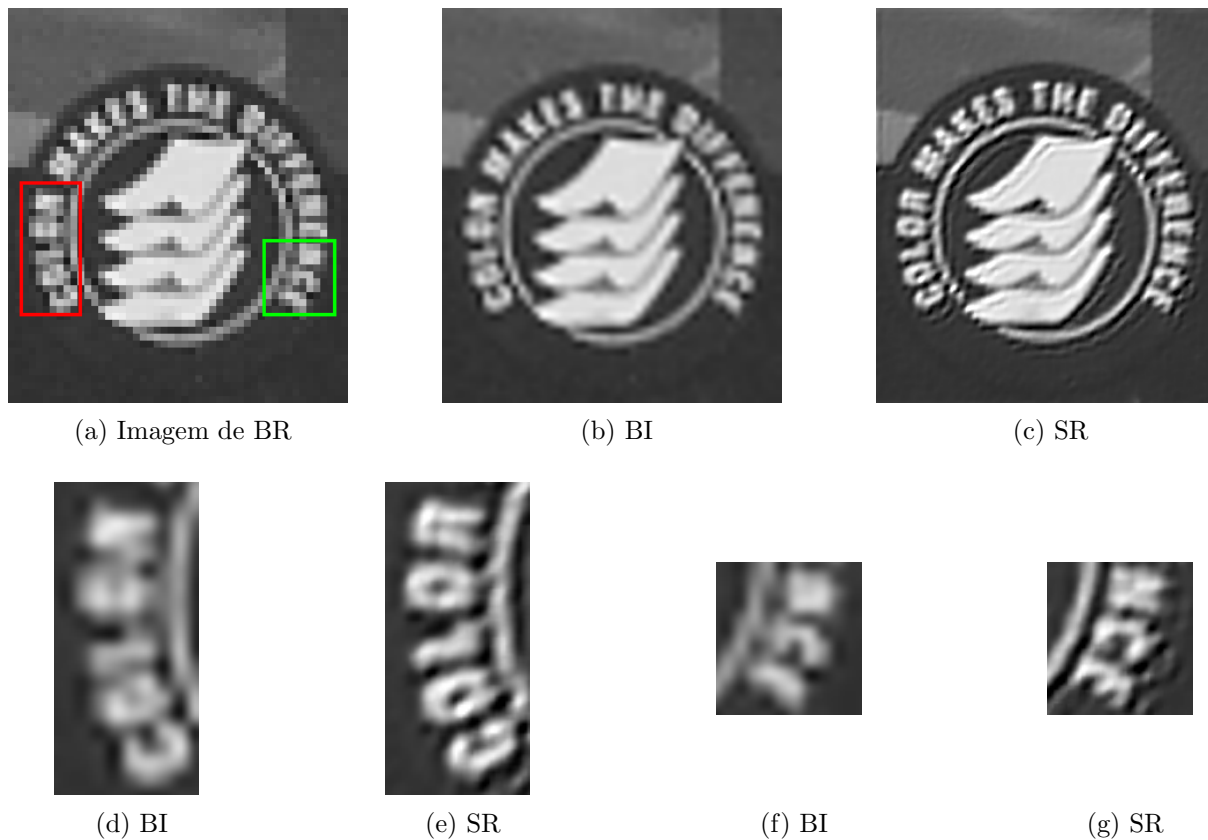


Figura 3.22: Resultado da SR para o conjunto B. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

ambos resultados (Figuras 3.26d e 3.26e), nota-se que a letra “A” não é legível no resultado obtido pelo método da interpolação Bicubica (Figura 3.26d). No entanto, no resultado da SR o formato da letra “A” tem boa definição.

O caso de teste G, foi realizado para o trabalho. Utilizando uma câmera térmica, várias imagens foram registradas consecutivamente de uma pessoa. A principal diferença notada nos resultados é com relação ao ruído. Pelo fato da técnica de interpolação Bicubica ter como referência uma única imagem o ruído presente na imagem de BR foi propagado para a hipótese de AR gerada (Figura 3.27b). No resultado da SR é possível notar que boa parte do ruído foi eliminado. Essa afirmação é melhor observada nos recortes realizados nos resultados. Nas Figuras 3.27d e 3.27e é possível perceber um “círculo” no plano de fundo da imagem. No resultado da SR, esse “círculo” apresenta bordas realçadas. A diminuição do efeito de ruído também auxilia na percepção do formato das dobras presentes na camisa (Figuras 3.27d e 3.27e). As partes mais escuras indicam que essa parte da camisa não está “colada ao corpo”, possuindo uma temperatura menor e conseqüentemente uma menor intensidade dos píxeis. Uma vantagem para aplicação da SR para imagens térmicas é a possibilidade de aferição da

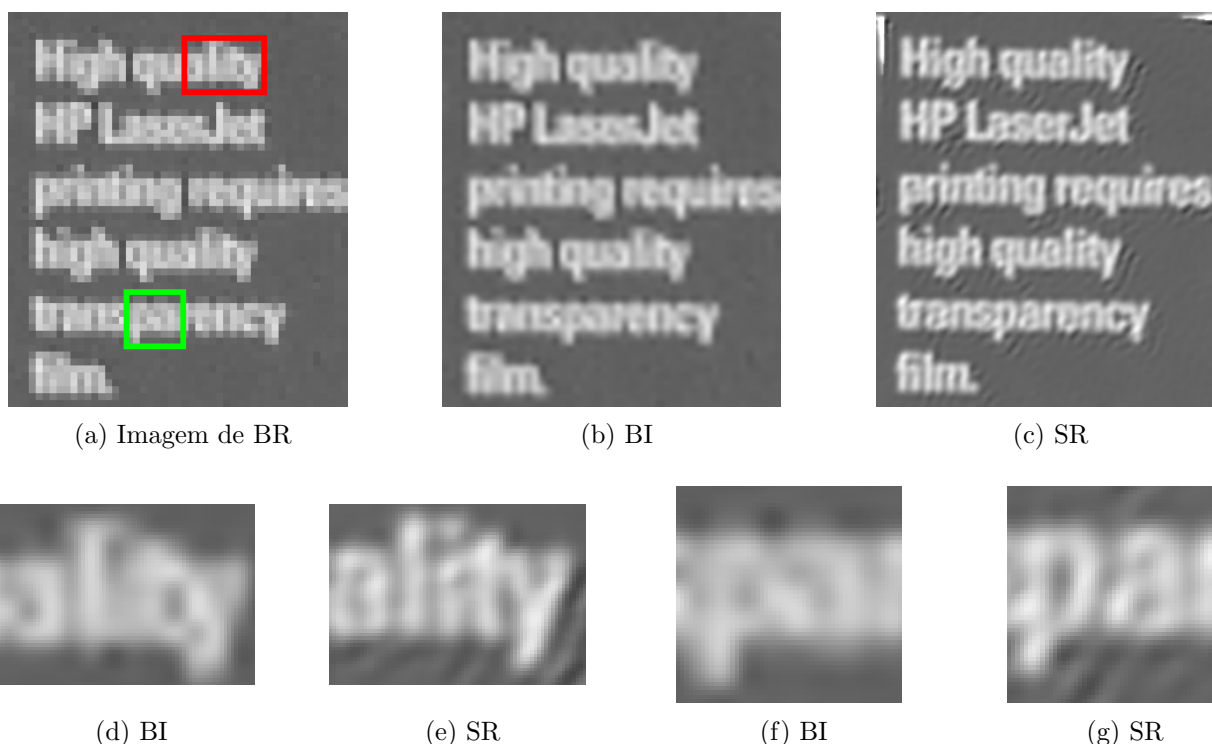


Figura 3.23: Resultado da SR para o conjunto C. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

temperatura pela intensidade do píxel. Com menos ruído, maior será a precisão na aferição.

3.4 Conclusão

Esse capítulo teve como primeiro objetivo a avaliação das técnicas de registro e reconstrução selecionadas pelo trabalho. Ao todo 4 métodos de registro e 5 métodos de reconstrução foram submetidas a vários testes, em um ambiente controlado, utilizando conjuntos de imagens de BR simuladas por meio de uma imagem de AR. As técnicas de registro e reconstrução foram experimentadas com o fim de encontrar a melhor combinação de técnicas para formar uma abordagem para aplicação da SR.

O método que sobressaiu na etapa de registro de imagens de BR foi o proposto por Keren (1988). O método se mostrou bastante eficiente para estimar as transformações de translação. No registro de rotação, o método teve um maior destaque, apresentando os melhores resultados para a minimização do erro dos movimentos de rotação. Além disso é o método com menor custo computacional.

Nos métodos de reconstrução o método de maior relevância foi o proposto por Maiseli (2014). Foram ao todo quatro métricas utilizadas para medir a qualidade da hipótese de AR gerada.

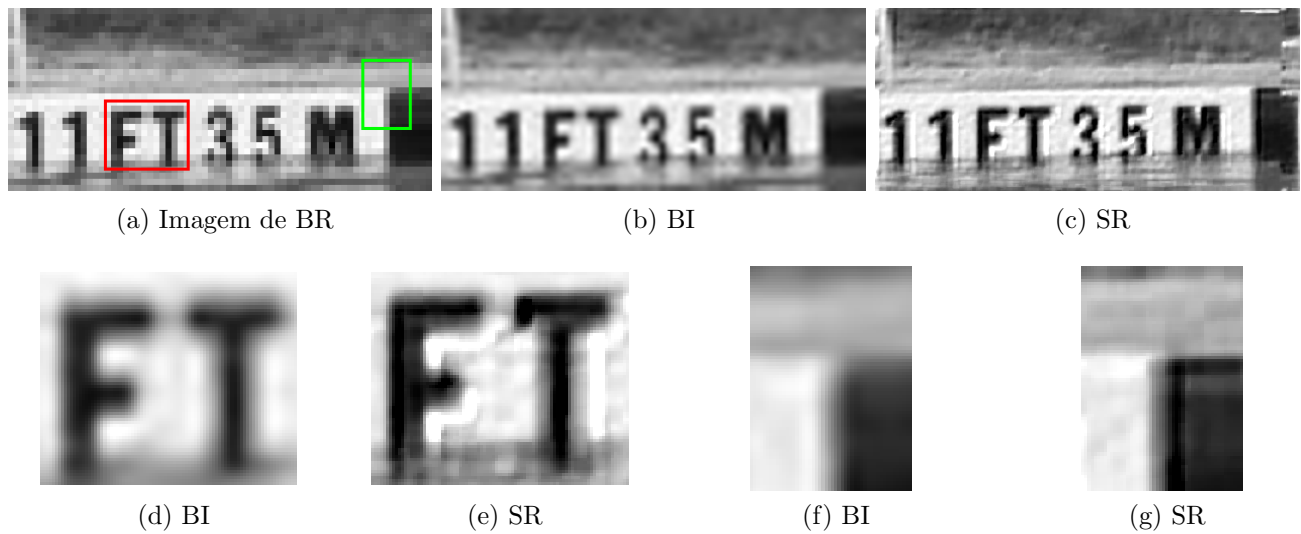


Figura 3.24: Resultado da SR para o conjunto D. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

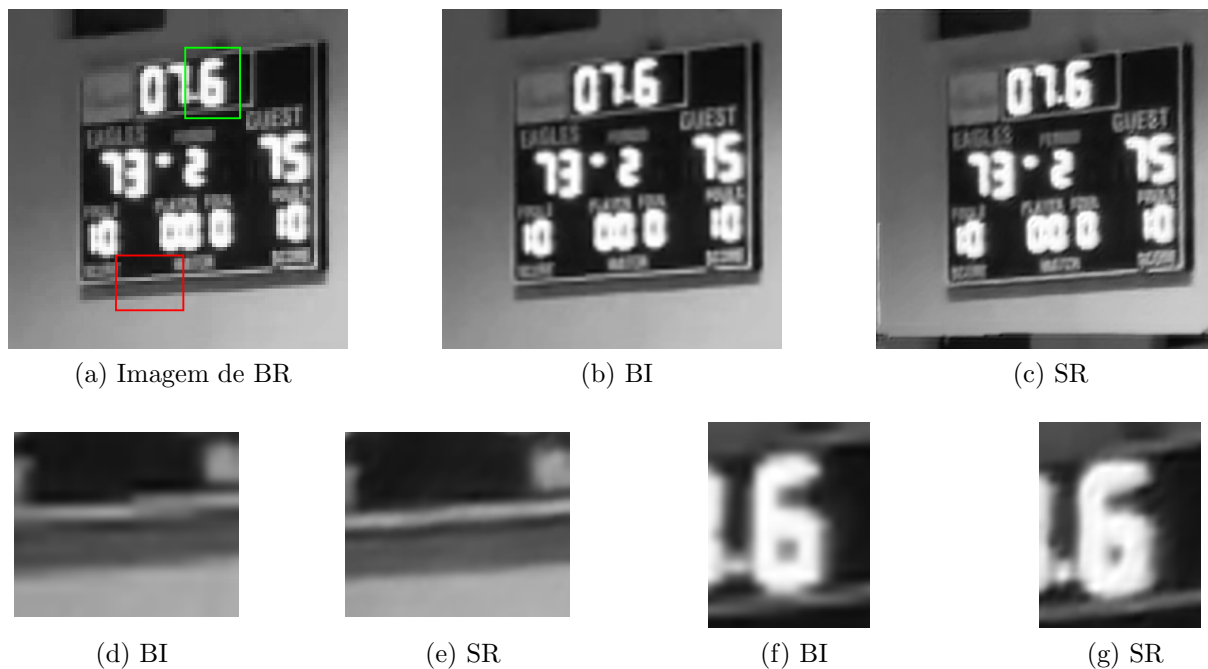


Figura 3.25: Resultado da SR para o conjunto E. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

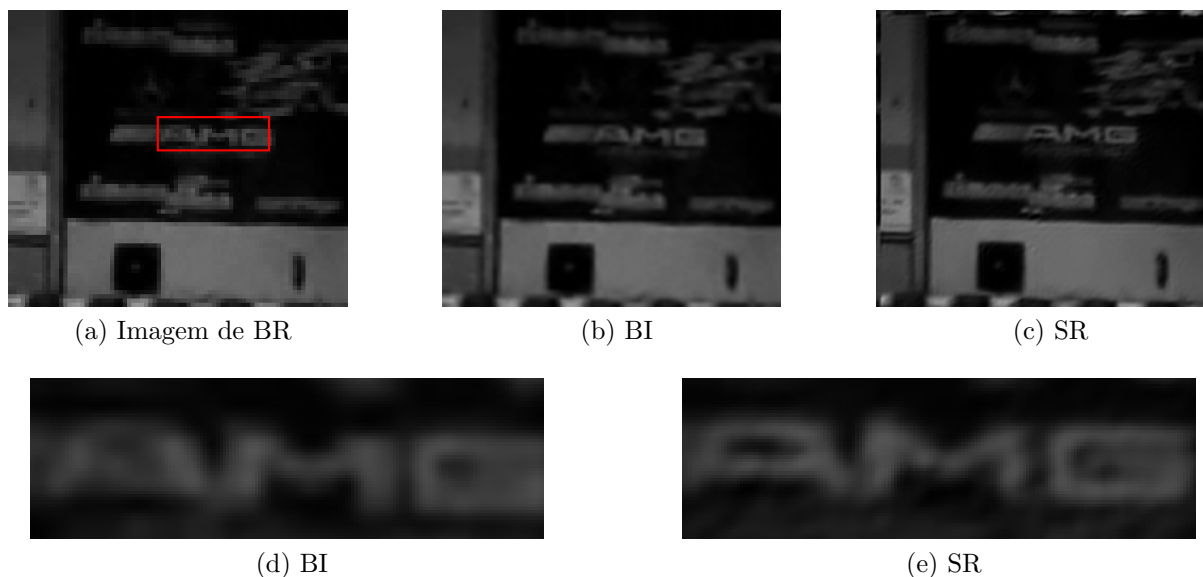


Figura 3.26: Resultado da SR para o conjunto F. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

O método obteve melhores resultados em três deles. Apesar de não apresentar um bom custo computacional, este foi escolhido pela melhor precisão em estimar a imagem AR, uma vez que os métodos avaliados não apresentaram resultados satisfatórios para problemas com FE alto.

Portanto, como resultado da fase de testes com imagens de BR simuladas, define-se, a combinação do método de Keren (1988) para a fase de registro e do método de Maiseli (2014) para a fase reconstrução, como a abordagem mais adequada para aplicação em problemas de SR.

Após a definição dessa combinação, um segundo objetivo foi proposto, validar a combinação escolhida. A abordagem de SR, pela combinação dos métodos escolhidos, foi aplicada a problemas com imagens reais para estimar a imagem de AR com um FE igual a 6. Porém neste caso, não se sabe quais são os deslocamentos de transformação entre as imagens e as degradações do processo de formação das imagens. Foram ao todo 7 conjuntos de imagens de BR utilizados nos experimentos, onde 6 foram obtidos em bancos de dados na literatura e um foi produzido para o trabalho.

Os resultados obtidos mostraram a validade e eficácia da escolha realizada para a abordagem de SR. Em todos os conjuntos avaliados, a SR teve resultados satisfatórios, evidenciando a capacidade de recuperar informações que não eram percebidas nas imagens de BR. As principais características a se ressaltar são, a boa recuperação das áreas de alta frequência (bordas) presentes na imagem e diminuição do efeito de ruído.

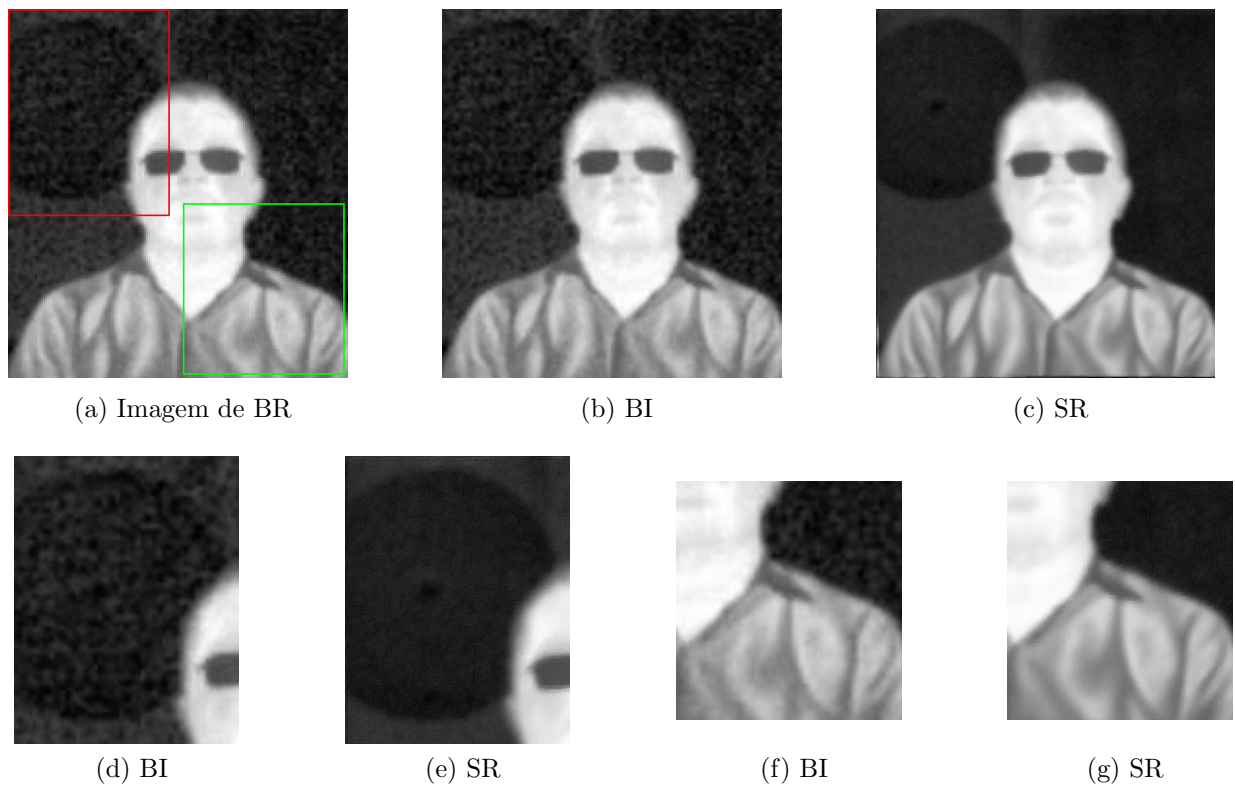


Figura 3.27: Resultado da SR para o conjunto G. (a) Uma imagem do conjunto de BR estendida, com as marcações de corte para análise; (b) Resultado da interpolação Bicubica; (c) Resultado da SR; (d) e (f) Corte no resultado da interpolação Bicubica; (e) e (g) Corte no resultado da SR.

Capítulo 4

Conclusão

Este trabalho apresentou um estudo de técnicas de registro e reconstrução de imagens, avaliando-as com o objetivo de obter a melhor combinação para sua aplicação em problemas de SR.

A SR é uma técnica que tem como objetivo gerar uma imagem de AR a partir de uma série de observações de BR de uma mesma cena. Algumas técnicas de aumento de resolução utilizam apenas uma imagem de BR para estimar a hipótese de AR. A SR, por sua vez, pode utilizar uma maior quantidade de imagens, adicionando mais informação ao problema. Para que o conjunto de imagens de BR seja relevante ao problema de SR é preciso que as imagens apresentem deslocamentos entre si a nível de sub-píxel.

São duas as principais etapas da SR, o registro e a reconstrução. Uma vez que a SR utiliza múltiplas imagens é necessário que, antes do processo de reconstrução as imagens de BR sejam registradas em um único plano de AR. E pelo fato dos deslocamentos serem em escalas mínimas é indispensável que os métodos de registro tenham uma alta precisão. Após o registro as imagens passam pela etapa de reconstrução. Os métodos de reconstrução são os responsáveis por estimar o processo de captura das imagens e tratar as possíveis degradações na imagem AR.

Quatro métodos de registro e cinco métodos de reconstrução foram avaliados. Os experimentos foram divididos em duas etapas, experimentos com imagens simuladas e experimentos com imagens reais. A primeira fase de experimentação teve como objetivo escolher o melhor método de registro e o melhor método de reconstrução, para que combinando-os, estes possam ser aplicados a problemas de SR. Ao todo foram testadas 128 imagens aplicadas a 6 casos de teste distintos.

Dentre os métodos de registro avaliados o que obteve melhor performance foi o método proposto por Keren (1988), apresentando as melhores médias nos três critérios avaliados, transformação de translação, rotação e tempo de execução. Com relação aos métodos de reconstrução, o método em destaque é o proposto por Maiseli (2014). O método teve as melhores médias nos critérios de avaliação de erro. Porém, o método não é eficiente quanto ao custo computacional. Este foi o método que teve as maiores médias de tempo de execução. Esses, portanto, foram os métodos escolhidos.

A segunda etapa de experimentos reuniu 7 conjuntos de imagens de BR com diferentes características, com o propósito de validar a escolha realizada. As imagens de AR produzidas

pela abordagem de SR foram comparados com a interpolação Bicubica. Os resultados obtidos foram considerados satisfatórios. Em 6 dos conjuntos de teste a SR, claramente, demonstrou superioridade. O método teve uma boa recuperação das áreas de alta frequência nas imagens, bem como a suavização do efeito do ruído das imagens de BR.

O trabalho mostrou a validade da aplicação da SR para problemas de aumento de resolução e a combinação de métodos escolhida trouxe resultados eficazes para as imagens avaliadas.

4.1 Trabalhos Futuros

Com relação ao registro de imagens de BR, sugere-se a investigação de outras técnicas que envolvam o registro de transformações diferentes das estudadas pelo trabalho, removendo a restrição da utilização de imagens com deslocamentos de translação e rotação apenas. Como pode ser visto no trabalho, a precisão do método de registro é essencial para o resultado da SR.

Finalmente, sugere-se a aplicação da abordagem de SR investigada no trabalho, através da combinação das técnicas de Keren (1988) e Maiseli (2014), a problemas práticos que envolvam imagens térmicas. O alto custo das câmeras térmicas de AR, muitas vezes, inviabiliza a sua aplicação. Através da técnica de SR, imagens térmicas registradas através de sensores de BR, podem ser utilizados para gerar a imagem de AR.

Referências Bibliográficas

- Adler, Jeremy e Parmryd, I. (2010). Quantifying colocalization by correlation: the pearson correlation coefficient is superior to the mander's overlap coefficient. *Cytometry Part A*, 77(8):733–742.
- Arroyo, J. E. C. (2002). *Heurísticas e meta heurísticas para otimização combinatória multiobjetivo*. PhD thesis, Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. Biblioteca Digital da Unicamp.
- Babacan, S Derin e Molina, R. e. K. A. K. (2011). Variational bayesian super resolution. *Image Processing, IEEE Transactions on*, 20(4):984–999.
- Baboulaz, Loïc e Dragotti, P. L. (2009). Exact feature extraction using finite rate of innovation principles with an application to image super-resolution. *Image Processing, IEEE Transactions on*, 18(2):281–298.
- Baker, Simon e Kanade, T. (2002). Limits on super-resolution and how to break them. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(9):1167–1183.
- Buades, Antoni e Coll, B. e. M. J.-M. (2005). A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530.
- Chai, Tianfeng e Draxler, R. R. (2014). Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7(3):1247–1250.
- Chan, Tony F e Osher, S. e. S.-J. (2001). The digital tv filter and nonlinear denoising. *Image Processing, IEEE Transactions on*, 10(2):231–241.
- Charbonnier, Pierre e Blanc-Féraud, L. e. A.-G. e. B. M. (1994). Two deterministic half-quadratic regularization algorithms for computed imaging. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 2, pages 168–172. IEEE.
- Chen, Ting e Catrysse, P. B. e. E. G. A. e. W. B. A. (2000). How small should pixel size be? In *Electronic Imaging*, pages 451–459. International Society for Optics and Photonics.
- Chikamatsu, Shingo e Nakaya, T. e. K. M. e. K. N. e. H. T. e. N. M. (2010). Super-resolution technique for thermography with dual-camera system. In *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pages 1895–1898. IEEE.

- Cossairt, O. (2011). *Tradeoffs and limits in computational imaging*. PhD thesis, Columbia University.
- Deshmukh, Manjusha e Bhosle, U. (2011). A survey of image registration. *INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)*, 5(3):245.
- Farsiu, Sina e Robinson, M. D. e. E. M. e. M. P. (2004). Fast and robust multiframe super resolution. *Image processing, IEEE Transactions on*, 13(10):1327–1344.
- Gao, Xinbo e Wang, Q. e. L. X. e. T. D. e. Z. K. (2011). Zernike-moment-based image super resolution. *Image Processing, IEEE Transactions on*, 20(10):2738–2747.
- Glasner, Daniel e Bagon, S. e. I. M. (2009). Super-resolution from a single image. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 349–356. IEEE.
- Hardie, R. (2007). A fast image super-resolution algorithm using an adaptive wiener filter. *Image Processing, IEEE Transactions on*, 16(12):2953–2964.
- Heinze, Norbert e Esswein, M. e. K. W. e. S. G. (2010). Image exploitation algorithms for reconnaissance and surveillance with uav. In *SPIE Defense, Security, and Sensing*, pages 76680U–76680U. International Society for Optics and Photonics.
- Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences*, 32(9):1283–1298.
- Irani, Michal e Peleg, S. (1991). Improving resolution by image registration. *CVGIP: Graphical models and image processing*, 53(3):231–239.
- Karch, Barry K e Hardie, R. C. (2015). Robust super-resolution by fusion of interpolated frames for color and grayscale images. *Frontiers in Physics*, 3:28.
- Keren, Danny e Peleg, S. e. B. R. (1988). Image sequence enhancement using sub-pixel displacements. In *Computer Vision e Pattern Recognition, 1988. Proceedings CVPR'88., Computer Society Conference on*, pages 742–746. IEEE.
- LCAV, A. C. L. (2015). Super-resolution. <http://lcav.epfl.ch/software/superresolution/index.html>.
- Lin, Frank C e Fookes, C. B. e. C. V. e. S. S. (2005). Investigation into optical flow super-resolution for surveillance applications. *APRS Workshop on Digital Image Computing: Pattern Recognition and Imaging for Medical Applications*.
- Lin, Zhouchen e Shum, H.-Y. (2004). Fundamental limits of reconstruction-based superresolution algorithms under local translation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(1):83–97.
- Lucchese, Luca e Cortelazzo, G. M. (2000). A noise-robust frequency domain technique for estimating planar roto-translations. *Signal Processing, IEEE Transactions on*, 48(6):1769–1786.

- Maiseli, Baraka e Elisha, O. e. M. J. e. G. H. (2014). Edge preservation image enlargement and enhancement method based on the adaptive perona–malik non-linear diffusion model. *IET Image Processing*, 8(12):753–760.
- Marcel, Baptiste e Briot, M. e. M. R. (1997). Calcul de translation et rotation par la transformation de fourier. *TS. Traitement du signal*, 14(2):135–149.
- MATLAB (2014). *Version 8.3.0.532 (R2014a)*. The MathWorks Inc., Natick, Massachusetts.
- Milanfar, P. (2010). *Super-resolution imaging*. CRC Press, New York.
- Milanfar, P e Farsiu, S. e. R. D. (2011). Mdsp super-resolution and demosaicing datasets. <https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html>.
- Morris, Nigel JW e Avidan, S. e. M. W. e. P. H. (2007). Statistics of infrared images. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–7. IEEE.
- Nasrollahi, Kamal e Moeslund, T. B. (2014). Super-resolution: a comprehensive survey. *Machine vision and applications*, 25(6):1423–1468.
- NZBRI, N. Z. B. R. I. (2015). Magnetic resonance imaging at 3 tesla - the first in new zealand.
- Pan, Zongxu e Yu, J. e. H. H. e. H. S. e. Z. A. e. M. H. e. S. W. (2013). Super-resolution based on compressive sensing and structural self-similarity for remote sensing images. *IEEE transactions on geoscience and remote sensing*, 51(9):4864–4876.
- Panagiotopoulou, Antigoni e Anastassopoulos, V. (2008). Super-resolution reconstruction of thermal infrared images. In *Proceedings of the 4th WSEAS International Conference on REMOTE SENSING*.
- Park, Sung Cheol e Park, M. K. e. K. M. G. (2003). Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36.
- Perona, Pietro e Shiota, T. e. M. J. (1994). Anisotropic diffusion. In *Geometry-driven diffusion in computer vision*, pages 73–92. Springer.
- Pham, Tuan Q e Van Vliet, L. J. e. S. K. (2006a). Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Applied Signal Processing*, 2006:236–236.
- Pham, T. Q. (2006b). Spatiotonal adaptivity in super-resolution of under-sampled image sequences. Master’s thesis, TU Delft, Delft University of Technology.
- Plenge, Esben e Poot, D. H. e. B. M. e. K. G. e. H. G. e. W. P. e. v. d. W. L. e. N. W. J. e. M. E. (2012). Super-resolution methods in mri: Can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time? *Magnetic Resonance in Medicine*, 68(6):1983–1993.

- Robinson, Dirk e Milanfar, P. (2004). Fundamental performance limits in image registration. *Image Processing, IEEE Transactions on*, 13(9):1185–1199.
- Robinson, Dirk e Farsiu, S. e. M. P. (2009). Optimal registration of aliased images using variable projection with applications to super-resolution. *The Computer Journal*, 52(1):31–42.
- Rudin, Leonid I e Osher, S. (1994). Total variation based image restoration with free local constraints. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 1, pages 31–35. IEEE.
- Sunkavalli, Kalyan e Joshi, N. e. K. S. B. e. C. M. F. e. P. H. (2012). Video snapshots: Creating high-quality images from video clips. *Visualization and Computer Graphics, IEEE Transactions on*, 18(11):1868–1879.
- Tian, Jing e Ma, K.-K. (2010). Stochastic super-resolution image reconstruction. *Journal of Visual Communication and Image Representation*, 21(3):232–244.
- Tsai, RY e Huang, T. S. (1984). Multiframe image restoration and registration. *Advances in computer vision and Image Processing*, 1(2):317–339.
- Vandewalle, Patrick e Sü, S. e. V. M. e. o. (2006). A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP Journal on Advances in Signal Processing*, 2006.
- Wallach, Daphné e Lamare, F. e. K. G. e. V. D. (2012). Super-resolution in respiratory synchronized positron emission tomography. *Medical Imaging, IEEE Transactions on*, 31(2):438–448.
- Wang, Yi e Fevig, R. e. S. R. R. (2008). Super-resolution mosaicking of uav surveillance video. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pages 345–348. IEEE.
- Weber, A. G. (1997). The usc-sipi image database: Version 5, original release: October 1997, signal and image processing institute, university of southern california, department of electrical engineering.
- Yang, Chih-Yuan e Ma, C. e. Y. M.-H. (2014). Single-image super-resolution: A benchmark. In *Computer Vision–ECCV 2014*, pages 372–386. Springer.
- Yang, Jianchao e Wright, J. e. H. T. S. e. M. Y. (2010). Image super-resolution via sparse representation. *Image Processing, IEEE Transactions on*, 19(11):2861–2873.
- Yoshida, Tomonari e Takahashi, T. e. D. D. e. I. I. e. M. H. (2012). Robust face super-resolution using free-form deformations for low-quality surveillance video. In *Multimedia and Expo (ICME), 2012 IEEE International Conference on*, pages 368–373. IEEE.
- Zhang, Xin e Lam, E. Y. e. W. E. X. e. W. K. K. (2008). Application of tikhonov regularization to super-resolution reconstruction of brain mri images. In *Medical Imaging and Informatics*, pages 51–56. Springer.

- Zhang, Liangpei e Zhang, H. e. S. H. e. L. P. (2010). A super-resolution reconstruction algorithm for surveillance images. *Signal Processing*, 90(3):848–859.
- Zhu, Xiao Xiang e Bamler, R. (2012). Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic sar. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(1):247–258.
- Zhu, Qiuxi e Li, X. e. M. W. (2013). Image super-resolution representation via image patches based on extreme learning machine. In *2013 International Conference on Software Engineering and Computer Science*. Atlantis Press.
- Zomet, Assaf e Rav-Acha, A. e. P. S. (2001). Robust super-resolution. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–645. IEEE.
- Zou, Wilman WW e Yuen, P. C. (2012). Very low resolution face recognition problem. *Image Processing, IEEE Transactions on*, 21(1):327–340.