

Análise Comparativa de Métodos de Fusão de Imagens

Eliamara S. da Silva, Carla L. Pagliari, Eduardo A. Barros da Silva e Andreas Ellmauthaler

Resumo— Em geral, dados oriundos de regiões espectrais diferentes aumentam a qualidade da informação para o observador. Este artigo aborda a fusão de imagens oriundas da região da luz visível e da região do infra-vermelho. Um ponto crítico da fusão é a seleção dos pontos 'relevantes', em cada uma das imagens, a serem combinados. Diversas técnicas de fusão são aplicadas às imagens fonte, onde são avaliados os efeitos causados pelas diferentes combinações de medidas de informação 'relevante' e outros parâmetros, incluindo métodos de decisão na imagem combinada. Os resultados são comparados, objetivamente e perceptualmente, a fim de se estipular combinações ótimas de parâmetros de fusão entre imagens de diferentes faixas de frequência. Uma nova combinação de regras de fusão, que produz ótimos resultados, é uma outra contribuição deste artigo.

Palavras-Chave— fusão de imagens, região do infra-vermelho, análise objetiva, análise subjetiva

Abstract— In general, data from different spectral regions convey additional quality information to the observer. This article studies fusion of infrared and visible light images. A key point to any fusion scheme lies in the selection of 'relevant' (feature) points, on each image, to be fused. Several fusion techniques are applied to the original images, in order to evaluate the effects inflicted by the combination of 'relevant' information measures, and other parameters, such as fusion decision methods to the fused image. The results are, objectively and subjectively, compared in order to find optimal parameter combinations for fusion of infrared and visible imagery. A new combination of fusion rules, that produces very good results, is another contribution of this work.

Keywords— image fusion, infrared region, objective analysis, subjective analysis

I. INTRODUÇÃO

O objetivo da fusão de imagens é combinar as informações oriundas de vários tipos de sensores (*multisensor* ou *multimodal*), de cenas adquiridas em diferentes períodos de tempo (*multitemporal*), de cenas capturadas com diferentes distâncias focais (*multifocus*), e de cenas capturadas de diferentes pontos de vista (*multiview*) em apenas uma imagem. Tal imagem deve apresentar um grau de 'qualidade' que não pode ser obtido, isoladamente, por cada uma das imagens fonte. Entende-se, neste caso, que o termo qualidade se refere ao tipo de informação que a imagem combinada apresenta, que varia com o tipo de aplicação [1].

Diversas aplicações que utilizam fusão de imagens de múltiplos sensores, em especial sensores nas regiões da luz visível e do infra-vermelho, estão presentes na área militar, na área médica, no resgate de vítimas de incêndios e desastres naturais, sensoriamento remoto, bem como nas áreas de vigilância. Este tipo de fusão enfatiza a informação de cada faixa de frequência. O desenvolvimento de algoritmos de fusão de imagens a partir de diferentes tipos de sensores também é

importante para aplicações na indústria, tais como inspeção de peças e equipamentos. Em particular nos casos em que um tipo de sensor não consegue exibir todas as informações desejáveis para a correta inspeção do objeto.

O objetivo da fusão multi-temporal é identificar mudanças ocorridas na mesma cena capturada em diferentes instantes de tempo. Este tipo de fusão é muito utilizado nas imagens médicas para detectar eventuais mudanças nos órgãos internos e tumores, além de ser aplicado na área de sensoriamento remoto para monitorar áreas de desmatamento.

A fusão multi-foco almeja ter como resultado uma imagem em que todos os objetos estejam em foco. Os algoritmos identificam as regiões em foco e as combinam em uma única imagem. No caso de fusão multi-vistas, um conjunto de imagens da mesma cena capturadas de diferentes pontos de vista são combinadas para se obter vistas virtuais [1]. Desta forma, o resultado final apresenta a visão de um ponto de vista que fisicamente não existe (sem ter sido adquirido por uma câmera real).

Em geral, dados oriundos de regiões espectrais diferentes aumentam a quantidade (e a qualidade) de informação para o usuário final. Sempre que tais imagens forem combinadas de maneira adequada, a imagem combinada (resultante) pode se transformar em uma importante ferramenta que vai oferecer mais informações do que duas imagens isoladas em suas respectivas faixas de frequência. O processo de fusão de duas imagens pode ser relativamente simples, como a obtenção da média ponderada de ambas as imagens por exemplo. No entanto, o objetivo é extrair as informações 'úteis' (relevantes para o tipo de aplicação) de cada uma das imagens, o que torna o processamento bem mais complicado.

A análise da qualidade das imagens resultantes de algoritmos de fusão é necessária. No entanto, esta tarefa é dificultada pela impossibilidade de comparação, no caso de fusão de imagens de regiões espectrais diferentes, com as imagens originais, inviabilizando – na maioria das vezes – o uso de métricas totalmente referenciadas. Além disto, o uso de apenas um tipo de medida pode não ser eficiente para analisar a qualidade da imagem combinada [2,3,4]. Desta forma, a inspeção visual é fundamental para avaliar a qualidade das imagens geradas pelos diferentes métodos de fusão [4].

Esta proposta emprega diferentes técnicas de fusão de imagens, variando as medidas de informação relevantes nas faixas da luz visível e do infra-vermelho a fim de estipular uma combinação ótima de parâmetros de fusão entre imagens destas diferentes faixas de frequência. Este artigo está organizado da seguinte forma. Na Seção II são apresentados os conceitos de fusão de imagens. As técnicas de fusão usadas são apresentadas na Seção III. A Seção IV exhibe, compara e analisa os resultados obtidos, e na Seção V são apresentadas as conclusões.

II. FUSÃO DE IMAGENS

Em geral, o processo de fusão é dividido em duas etapas: o registro da imagem (*image registration*), que mapeia as imagens em um único sistema de coordenadas, e a combinação de atributos das imagens (e.g. intensidade dos *pixels*, cores, etc) [1]. Este artigo aborda a fusão de imagens obtidas de dois tipos de sensores. Neste caso, a captura das imagens ocorre em diferentes comprimentos de onda usando-se sensores que

Eliamara S. da Silva e Carla L. Pagliari, Departamento de Engenharia Elétrica, Instituto Militar de Engenharia, Rio de Janeiro, Brasil; Eduardo A. Barros da Silva e Andreas Ellmauthaler, Programa de Pós Graduação em Engenharia Elétrica, COPPE/UFRJ, Rio de Janeiro, Brasil Este trabalho foi financiado pelo programa CAPES/Pró-Defesa

podem apresentar resoluções espaciais diferentes, faixas dinâmicas diferentes e campos de vista (field-of-view, FOV) diferentes.

Um dos primeiros passos da fusão é o processo de registro (*registration*) de ambas imagens. Ou seja, o mapeamento das duas imagens em um único sistema de coordenadas. Quando dois conjuntos de dados, no caso *pixels*, são amostrados em perspectivas diferentes, diz-se que encontram-se em sistemas de coordenadas diferentes [5]. O processo de registro, é o processo de se transformar os dois conjuntos de *pixels* em apenas um sistema de coordenadas. E, tal processo torna-se ainda mais difícil quando as imagens pertencem a diferentes regiões espectrais, uma vez que o processo de registro necessita identificar os pontos homólogos (correspondentes) nas duas imagens. Este processo de casamento (*matching*) já é bastante complexo quando ambas as imagens pertencem à mesma faixa de frequência [5]. Tal processamento é fundamental para que a fusão das imagens seja bem sucedida. Notar que somente depois de ter sido realizado o casamento (estabelecimento das correspondências espaciais) é que pode-se garantir que as informações dos diferentes sensores estão espacialmente alinhadas e podem ser combinadas. Caso contrário, o algoritmo pode combinar áreas diferentes da imagem.

Um exemplo de fusão de imagens é dado pela Fig 1. A Fig 1(a) mostra a cena na região da luz visível, a Fig 1(b) exibe a mesma cena na região do infra-vermelho (IV), e a Fig 1(c) mostra o resultado da combinação linear das duas. Nota-se que o homem, antes escondido pela vegetação na Fig 1(a), aparece na Fig 1(c). O mesmo ocorre com detalhes antes não visíveis na Fig 1(b), como parte do telhado da casa e a cerca que passam a estar presentes da Fig 1(c). Ou seja, as informações ditas 'relevantes' aparecem na imagem combinada. No entanto, a imagem do homem, antes representada com um nível de cinza mais claro, aparece em um tom mais escuro após o processo de fusão. A imagem fundida apresenta um baixo nível de contraste global. Neste caso, ainda é possível perceber que existe uma pessoa na cena. Mas, dependendo da imagem, um método de fusão tão simples pode não produzir bons resultados.

Um dos problema da fusão de imagens é o processo de seleção dos pontos relevantes (que dependem do tipo de aplicação) das imagens originais e a sua representação em apenas uma única imagem. Tais algoritmos tem como objetivo criar uma imagem que contenha as informações relevantes de cada imagem fonte sem, na medida do possível, introduzir artefatos. A seguir são apresentados alguns métodos de fusão de imagens, bem como os utilizados nas simulações.

III TÉCNICAS DE FUSÃO DE IMAGENS

A combinação linear ponderada, representada em (1), é de fácil implementação e rápida execução:

$$F(x, y) = w_A A(x, y) + w_B B(x, y) \quad (1)$$

onde $F(x,y)$ é a imagem combinada, w_A e w_B são escalares e $A(x,y)$ e $B(x,y)$ são as intensidades dos *pixels* nas coordenadas (x,y) .

No entanto, apesar da baixa complexidade computacional, as imagens resultantes do processo de fusão apresentam baixo contraste e baixo nível de detalhes (Fig 1(c)). Estes problemas podem ser reduzidos a partir de uma seleção 'ótima' dos pesos w_A e w_B , usando a técnica denominada PCA (Principal Component Analysis). Este técnica determina, usando a matriz de covariância para as duas imagens, os valores dos pesos de forma a maximizar a variância do valor da intensidade na imagem fundida. O método usa como medida de relevância (ou de interesse) a variância global, sempre estipulando um peso maior para a imagem fonte que apresenta a variância mais alta. Este fator deixa o método suscetível a ruído, a '*dead pixels*'

(ocorrência muito comum em sensores na região do infra-vermelho), e outros artefatos indesejáveis que aumentam a variância total da imagem fonte.

Estudos mostram que o Sistema Visual Humano (SVH) contém 'canais sintonizados' em um conjunto de frequências espaciais e orientações (detalhes verticais, horizontais, etc). Uma cena do mundo real pode conter objetos de vários tamanhos e estes objetos apresentam pontos de interesse (detalhes) de diferentes tamanhos. Além disto, os objetos podem estar localizados a distâncias distintas do observador. Assim, qualquer método que seja aplicado a apenas uma escala (espacial) pode estar 'perdendo' informações contidas em outras escalas. O método piramidal representa as informações localizadas no espaço (diferentes escalas) e na frequência. O esquema da imagem piramidal consiste em uma série de cópias filtradas por filtros passa-baixas e passa-altas da mesma imagem em diferentes escalas [4].

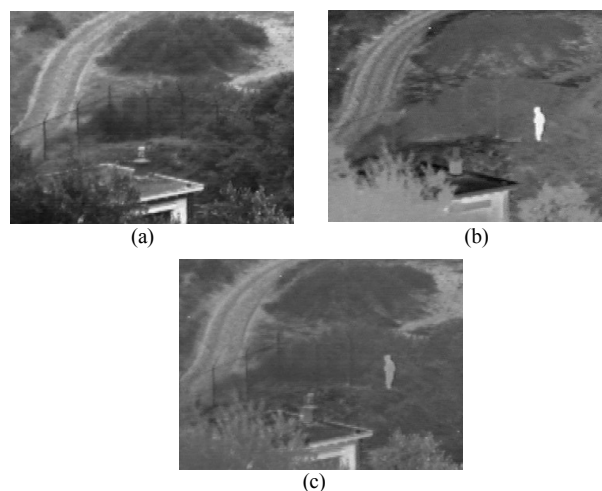


Fig 1. UN CAMP, frame 16 (a) visível; (b) IV; (c) resultado da fusão.

Os métodos de multi-resolução são de alta complexidade computacional e tendem a apresentar variações locais de contraste (bordas) nas diferentes resoluções e orientações. Este tipo de algoritmo produz imagens com contraste elevado, com as bordas preservadas e apresentando mais informações relevantes de ambas as imagens fonte. Com o SVH é sensível a variações locais de contraste (bordas), que representam transições rápidas na imagem e, em geral, delimitam os objetos, um método de fusão que use a informação de bordas é bastante desejável. Os métodos de fusão que usam algoritmos de multi-resolução permitem que pontos característicos das imagens sejam tratados em diferentes escalas e orientações. Além disto, nas bordas os valores dos coeficientes (e.g. da *wavelet*) são, em geral, elevados.

A transformada *wavelet* introduzida por Mallat [6], é uma ferramenta de análise que fornece uma decomposição multi-resolução de uma imagem em bases bi-ortogonais e resultados em uma representação não redundante da imagem. Essas bases são chamadas *wavelets*, e são funções geradas de uma única função chamada *wavelet* mãe pelas operações de dilatação e translação. No caso na transformada *wavelet* bi-dimensional é acrescentada a operação de rotação. A transformada *wavelet* decompõe a imagem em sub-bandas com diferentes informações de conteúdo em frequência (detalhes). São criadas, a cada nível de decomposição, quatro novas sub-bandas com metade da resolução espacial e escala de entrada da pirâmide. Cada nível contém uma versão da imagem original filtrada por um filtro passa-baixas, denominada imagem de aproximação, e três imagens filtradas por filtros passa-altas (detalhes verticais, horizontais e diagonais) da imagem original, denominadas sub-imagens (imagens de

detalhes) [6, 7]. Os métodos de multi-resolução, em particular os que usam a transformada *wavelet* são os que produzem os melhores resultados no processo de fusão de imagens [8].

A. Métodos de multi-resolução ('nível de pixel')

A Fig 2 ilustra o processo de fusão de imagens usando técnicas de multi-resolução (MR), no caso a transformada *wavelet* [7]. O processo aplica a transformada (DWT) nas imagens originais (x_A e x_B), seguida de alguma regra de fusão (representada pelos blocos denominados similaridade e atividade), resultando em uma representação desta combinação no domínio da transformada (y_F). A imagem fundida, x_F , é obtida aplicando-se a transformada inversa (IDWT) na representação desta combinação. Onde y_A e y_B , são os coeficientes da *wavelet* das imagens A e B, respectivamente, no domínio da transformada. A medida de atividade, que denota o nível de relevância de cada coeficiente, e a de similaridade, que quantifica o grau de correlação entre os coeficientes das duas imagens, populam os mapas de decisão e fusão. O mapas de decisão e fusão determinam a combinação dos coeficientes da *wavelet*. Como cada coeficiente da *wavelet* (ou um conjunto de coeficientes dentro de uma janela) é processado um a um (*pixel-a-pixel*), o método é classificado como sendo a 'nível de *pixel*'. A seguir, é apresentada uma análise mais detalhada do processo.

O bloco da transformada *wavelet* (DWT) calcula as decomposições multi-resoluções das imagens fontes. Pode-se definir o tipo da decomposição (*wavelets* ou *wavelets* complexas), os filtros utilizados e o número de níveis de decomposição. Uma grande parte da pesquisa em fusão de imagens tem dado ênfase no processo de escolha da representação MR que facilite a seleção e a combinação de regiões de interesse relevantes para a aplicação desejada [8].

Seja $p = (m; n; k; l)$ o índice do coeficiente da *wavelet*, onde m,n indicam a posição espacial em uma dada banda de frequência (orientação), k indica o nível de decomposição, e l a banda de frequência.

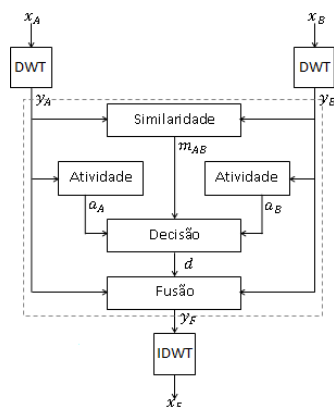


Fig 2. Processo de fusão de imagens ('nível de *pixel*').

Para o cálculo do grau de atividade (a_A e a_B) dos coeficientes da *wavelet* $y_A(p)$ e $y_B(p)$ de cada imagem, ou seja, o nível de importância (ou relevância) para uma particular aplicação de fusão de imagens, foram consideradas as seguintes medidas que operam em cada nível k , em cada banda de orientação l , e localização m,n :

- Valor Absoluto (VA) de cada coeficiente;
- Valor Absoluto (VA-Janela), em uma janela 3 x 3 centrada em cada coeficiente, a medida de atividade é

dada pela soma do valor absoluto de todos os coeficientes dentro dessa janela;

- Frequência espacial (FE) da sub-imagem;
- Filtro de Classificação (FC), em uma janela 3 x 3 centrada em cada coeficiente, a medida de atividade é dada pelo o coeficiente de maior valor absoluto dentro dessa janela;
- Filtro de mediana (FM), para um tamanho de janela 3 x 3, é obtido o valor do filtro de mediana para cada coeficiente.

A medida de similaridade (m_{AB}), em cada nível k , em cada banda de orientação l , e localização (m,n) , é determinada com o propósito de quantificar o grau de correlação entre as imagens fonte. A correspondência ou similaridade entre os coeficientes das transformadas das imagens originais são expressos em termos de medida de correlação local sobre a vizinhança das amostras. Analisando a medida de similaridade, podemos determinar o quanto as fontes diferem, usando tal resultado para ponderar a decisão [7].

O bloco denominado decisão, é o ponto chave para a fusão das imagens, uma vez que determina como os coeficientes serão combinados pelo processo de fusão. As opções consideradas neste trabalho são apresentadas a seguir [7]:

- Média dos coeficientes (M), o valor de y_F (coeficiente da *wavelet*) é dado pela média dos coeficientes y_A e y_B ;
- Máxima seleção (MS), o valor de y_F é dado pelo maior valor de medida de atividade de cada sub-imagem;
- Máxima seleção (MSP), semelhante a opção anterior, sendo que a medida de atividade de todas as sub-bandas de cada nível k são somadas e consideradas como uma única medida de atividade, ou seja, aplica-se a mesma regra de decisão para todas as sub-bandas;
- Combinação ponderada (CPMS), o valor de y_F é obtido por uma combinação ponderada em que os pesos são determinados com base na medida de similaridade;
- Combinação ponderada (CPMP) o valor de y_F é dado pela combinação ponderada de y_A e y_B em que os pesos de ponderação são os valores das suas respectivas medidas de atividade.

O módulo de fusão realiza a combinação dos coeficientes *wavelet* das imagens fonte de acordo com as regras de fusão acima estabelecidas.

B. Métodos de multi-resolução ('nível de região')

Este tipo de método visa combinar regiões e não somente *pixels*. Uma das vantagens é que o processo de fusão fica menos suscetível ao ruído e produz menos efeitos de *blurring*. Desta forma, pode-se identificar pontos de interesse nas imagens, e aglutiná-los em regiões usando algoritmos de segmentação. A Figura 3 mostra o processo de fusão de imagens a 'nível de região'. Este processo é uma extensão do processo de fusão de imagens do tipo multi-resolução a 'nível de *pixel*'. Uma versão deste tipo de método é apresentado em [7].

A principal diferença consiste na inclusão de um bloco de segmentação que usa as duas imagens fonte para obter um único mapa de segmentação (\mathcal{R}), que representa as regiões afim de ambas as imagens. Um algoritmo de segmentação é aplicado às imagens x_A e x_B , e combinado de forma a gerar um único mapa de segmentação em diferentes níveis de

decomposição, k . O algoritmo de segmentação é com base no algoritmo de Maximização de Expectativas [9]. As medidas de atividade e similaridade são computadas para cada região a cada nível de decomposição, k . Estas medidas são integradas produzindo um mapa de decisão, d , que por sua vez vai determinar a combinação (fusão) dos coeficientes das imagens fonte, resultando na imagem x_F .

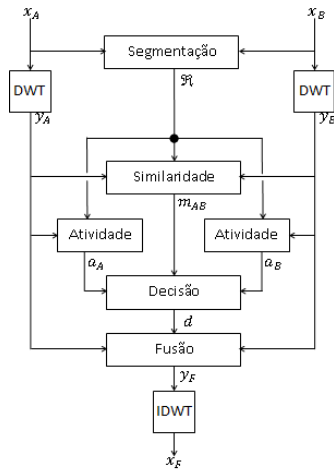


Fig 3. Processo de fusão de imagens('nível de região').

IV ANÁLISE DOS RESULTADOS

A seguir, são apresentados resultados selecionados das simulações. Um dos problemas enfrentados por este trabalho de pesquisa foi a obtenção de imagens (vídeos) nas duas regiões espectrais. Nas simulações foram utilizados o frame 16 dos vídeos UN CAMP(Fig 1), e o frame 21 dos vídeos OCTEC (Fig. 4) nas regiões da luz visível e do infra-vermelho [10].

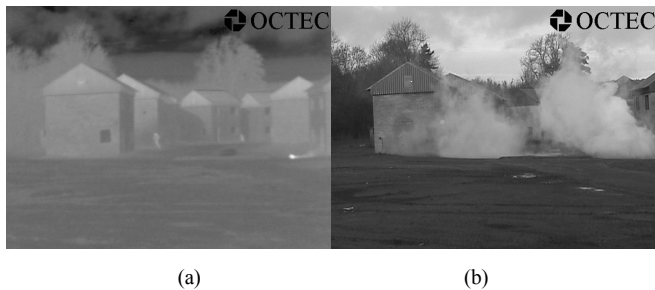


Fig 4. OCTEC, frame 21 (a) visível; (b) IV.

Avaliação objetiva. As métricas utilizadas para análise dos resultados foram: desvio padrão, entropia, gradiente médio, informação mútua, métrica com base no índice de qualidade da imagem (Qabf) proposta em [11] e a métrica (Qp) introduzida em [2] que mede a quantidade de informação transferida das imagens fonte para a imagem fundida.

Um problema da análise objetiva é saber se as imagens fundidas conseguem exibir as informações relevantes contidas nas duas imagens. Além das análises objetivas, são apresentados os resultados das análises subjetivas, uma vez que a avaliação perceptual permanece como o método mais confiável de análise da qualidade da fusão de duas imagens.

Avaliação subjetiva. Os testes exibiram as imagens fonte para os observadores, e em seguida as imagens fundidas com os diferentes combinações de parâmetros. Os resultados do MOS (Mean Opinion Score) são apresentados nas Tabelas 3 e 5. Cada observador associou uma nota variando de 5 a 1, correspondendo à qualidade (excelente, boa, regular, pobre, ruim) de cada imagem fundida.

As imagens fundidas foram geradas a partir de várias

combinações de regras de fusão, diferentes variações das medidas de atividade e decisão apresentadas. Todas as combinações foram consideradas para os processos de fusão multi-resolução a 'nível de pixel' (FP) e para os processos de fusão a 'nível de região' (FR), usando métodos de MR. Um terceiro tipo de combinação de fusão (FRP) é proposto neste artigo, onde as imagens aproximadas são processadas 'nível de região' e as imagens de detalhe (sub-imagens) a 'nível de pixel'.

Foram efetuadas simulações com 2, 4, 6 e 8 níveis de decomposição utilizando a DWT (Discrete Wavelet Transform) e a DT-CWT (Dual-Tree Complex Wavelet Transform) [12]. A partir dos resultados, observamos que o número de níveis de decomposição pode influenciar no desempenho da fusão de imagens. Entretanto, o uso de mais níveis de decomposição não produz resultados tão superiores que justifiquem o uso de mais níveis de decomposição. As simulações foram realizadas usando 4 níveis de decomposição por terem apresentado o melhores resultados para as imagens testadas.

Diversos tipos de filtro foram implementados para todos os casos com o propósito de analisar a influência dos filtros. Imagens combinadas com bordas pouco definidas foi um dos efeitos notados da aplicação de diferentes filtros. Os que apresentaram os melhores resultados foram o filtro biortogonal [6] para DWT e o filtro simétrico próximo [12] para DT-CWT.

Testes subjetivos e objetivos foram empregados na análise dos resultados das imagens fundidas para todas as 8 (oito) combinações para a seleção das melhores medidas de atividade, considerando os três tipos de fusão: FP, FR e FRP. O critério de decisão utilizado é o de Máxima Seleção (MS) (seção III-A) para todos os casos. Assim, de acordo com os testes subjetivos, as 3 (três) melhores combinações utilizando diferentes medidas de atividade foram selecionadas, e são apresentadas nas Tabelas 1 e 2, para as imagens OCTEC e UN CAMP respectivamente. A DT-CWT foi a transformada escolhida, por apresentar os melhores resultados nas diversas combinações. Os tipos de fusão selecionados pelos testes subjetivos foram o FP e o FRP.

TABELA 1. Combinações de Parâmetros para seleção de Medidas de Atividade - imagem OCTEC

	Método de MR	Tipo de Fusão	Medida de Atividade	
			Passa Baixas	Passa Altas
Combinação 1	DT-CWT	FRP	FE	FE
Combinação 2	DT-CWT	FRP	FM	FC
Combinação 3	DT-CWT	FRP	FM	VA

TABELA 2. Combinações de Parâmetros para seleção de Medidas de Atividade - imagem UN CAMP

	Método de MR	Tipo de Fusão	Medida de Atividade	
			Passa Baixas	Passa Altas
Combinação 1	DT-CWT	FP	FE	FE
Combinação 2	DT-CWT	FRP	FM	FC
Combinação 3	DT-CWT	FRP	FM	VA

Em seguida, 20 (vinte) combinações de métodos de decisão foram testadas em conjunto com os parâmetros descritos nas Tabelas 1 e 2. A Tabela 3 apresenta a melhor combinação para a imagem OCTEC de acordo com os testes subjetivos. A partir dos resultados da avaliação perceptual para a mesma imagem, foram selecionados os valores correspondentes das métricas objetivas, apresentados na Tabela 4. É importante ressaltar que tais valores não são os melhores de acordo com cada métrica. Neste caso, não houve

uma correlação direta entre os resultados das avaliações objetiva e subjetiva. Ou seja, nenhuma das medidas objetivas foi capaz de associar as melhores notas aos resultados escolhidos pelos observadores humanos.

TABELA 3. Melhor combinação de parâmetros para a imagem OCTEC de acordo com a avaliação subjetiva

Método de MR	Tipo de Fusão	Medida de Atividade		Decisão		MOS
		Passa Baixas	Passa Altas	Passa Baixas	Passa Altas	
DT-CWT	FRP	FM	FC	MS	MS	4,67

TABELA 4. Resultado da avaliação objetiva para a imagem OCTEC

MÉTRICAS DE AVALIAÇÃO OBJETIVA					
Desvio Padrão	Entropia	Qabf	Gradiente Médio	Inf. Mútua	Qp
58.6322	7,1764	0,6181	3,7379	0,4689	0,8820

A Tabela 5 apresenta o conjunto de parâmetros (métodos de MR e medidas) que combinados resultaram nas melhores imagens fundidas de acordo com as análises subjetivas para a imagem UN CAMP. Enquanto que os resultados da avaliação objetiva estão apresentados na Tabela 6. Desta vez, as melhores medidas objetivas tem associação com os resultados escolhidos pelos observadores humanos. Embora nenhuma delas tenha se destacado.

TABELA 5. Melhores combinações de parâmetros para a imagem UN CAMP de acordo com a avaliação subjetiva

Combinação	Método de MR	Tipo de Fusão	Medida de Atividade		Decisão		MOS
			Passa Baixas	Passa Altas	Passa Baixas	Passa Altas	
1	DT-CWT	FRP	FM	FC	CPMS	MS	3,67
2	DT-CWT	FRP	FM	VA	CPMS	MS	3,67
3	DT-CWT	FP	FE	--	CPMP	M	3,67
4	DT-CWT	FRP	FM	--	CPMP	M	3,67
5	DT-CWT	FRP	FM	FC	CPMP	CPMP	3,67
6	DT-CWT	FRP	FM	VA	CPMP	CPMP	3,67

TABELA 6. Resultados da avaliação objetiva para a imagem UN CAMP

Combinação	MÉTRICAS DE AVALIAÇÃO OBJETIVA					
	Desvio Padrão	Entropia	Qabf	Gradiente Médio	Inf. Mútua	Qp
1	32,4044	6,8947	0,4526	4,5881	0,1512	0,6777
2	32,6531	6,9149	0,4564	4,8904	0,1466	0,6785
3	26,5331	6,5445	0,3499	2,9489	0,1166	0,4985
4	24,4713	6,3065	0,3489	2,8766	0,1208	0,4939
5	25,6833	6,3634	0,4254	3,4003	0,1226	0,6379
6	26,2419	6,4177	0,4501	3,8120	0,1202	0,6625

A imagem resultante para a combinação de algoritmos da Tabela 3 das imagens fonte OCTEC nas regiões da luz visível e do infra-vermelho é apresentada na Fig. 5. Analisando os valores da Tabela 6, a Combinação 2 apresenta altos valores de entropia, os melhores resultados para as métricas Qabf e Qp, e bons resultados nas outras métricas. A imagem resultante desta combinação é apresentada na Fig.6.

V CONCLUSÕES

A escolha das regras de fusão depende das imagens fonte, como observamos nas combinações 'ótimas' para as duas imagens OCTEC e UN CAMP, em que os melhores resultados obtidos utilizam combinações diferentes de algoritmos. Uma comparação, por inspeção visual, pode ser realizada entre as Fig 6 e Fig 1(c), onde o resultado exibido pelo Fig 6 é claramente superior ao mostrado pelo Fig 1(c). Resultados

adicionais encontram-se disponíveis em <http://lpi.ime.br/fusao>.



Fig 5. Imagem resultante da melhor Combinação da Tabela 4 – OCTEC



Fig 6. Imagem resultante da Combinação 2 da Tabela 6 - UN CAMP

O tipo de fusão híbrida (FRP), proposto neste artigo, é o que produz os melhores resultados, assim como o uso da DT-CWT. A medida de atividade FM foi utilizada mais vezes, embora não se possa afirmar que seja a opção ótima.

A avaliação dos resultados dos algoritmos de fusão de imagens ainda é um problema crítico. As métricas objetivas tentam reproduzir as avaliações de um observador humano, apesar de apresentarem bons resultados, ainda não tem a capacidade de reproduzir tais avaliações completamente, nem sempre os melhores valores das métricas objetivas significam os melhores resultados selecionados pela análise subjetiva.

AGRADECIMENTOS

Os autores agradecem ao programa CAPES/Pró-Defesa que financia as pesquisas na área de fusão de imagens.

REFERÊNCIAS

[1] J. Flusser, F. Sroubek e B. Zitová, "Image Fusion: Principles, Methods, and Applications", Tutorial da 15ª Conferência Europeia de Processamento de Sinal (EUSIPCO 2007).

[2] C. S. Xydeas e V. Petrovic, "Objective image fusion performance measure", *Electronics Letters*, 38, pp.308-309, 2002.

[3] Z. Wang e A. C. Bovik, "A universal image quality index", *IEEE Signal Processing Letters*, vol 9, no. 3, pp. 81-84, March 2002.

[4] F. Sadjadi, "Comparative Image Fusion Analysis", *IEEE CVPR'05*, 2005.

[5] B. Zitová e J. Flusser, "Image registration methods: a survey", *Image and Vision Computing* 21, pp. 977-1000, 2003.

[6] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Anal. Machine Intell.*, 11, pp. 674-693, 1989.

[7] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions", *Information Fusion*, pp 259-280, 2003.

[8] M. Ghanous, S. Ghosh, M. Bayoumi, "A gradient-based hybrid image fusion scheme using object extraction", *IEEE International Conference on Image Processing 2008, ICIP2008*, pp.1300-1303.

[9] T. K. Moon, "The Expectation Maximization Algorithm", *IEEE Signal Processing Magazine*, Novembro 1996.

[10] <http://www.imagefusion.org/>

[11] G. Piella, H. Heijmans, "A new quality metric for image fusion", *IEEE International Conference on Image Processing 2003, vol III*, pp 173-176

[12] N G Kingsbury, "A Dual-Tree Complex Wavelet Transform with improved orthogonality and symmetry properties", *IEEE International Conference on Image Processing, ICIP2000*.