

Estudo comparativo de seis métodos para *Super Resolution* 2 versões para *Pocs*, *Iterated Back Projection*, *Robust Super Resolution*, *Structure-Adaptive Normalized Convolution* e composição de imagens realçadas.

Daniel da Silva Diogo Lara
daniels@dcc.ufmg.br

Programa de Pós-Graduação em Ciência da Computação – UFMG
Novembro – 2008

Resumo

Em muitas aplicações eletrônicas, imagens de alta resolução são necessárias (*e.g* Aplicações médicas, imagens de satélites, etc.). Super-Resolução é o termo dado a algoritmos que aumentam o poder de resolução da imagem a partir de várias outras imagens de baixa resolução da mesma cena. Esse trabalho apresenta uma análise comparativa dos principais métodos de super-resolução encontrados na literatura e ainda propõe um algoritmo eficiente para obtenção de uma imagem de mais alta resolução.

1. Introdução

A resolução da imagem é um dos parâmetros que limitam o projeto de câmeras digitais. Em geral, câmeras de mais alta resolução apresentam uma imagem de melhor qualidade visual permitindo enxergar mais detalhes.

Desta maneira, métodos eficientes para o aumento da resolução de imagens representam uma ferramenta importante para a melhoria de equipamentos de aquisição de imagens digitais.

Alem disso, métodos para o aumento da resolução de imagens podem ser úteis para melhoria de imagens antigas já adquiridas com resolução mais baixas.

Ainda existem categorias de imagens em que a *super resolution* não é necessária mas sim desejada, como exemplo, pode-se citar as aplicações amadoras pessoais.

Um aumento da resolução pode ser obtido por meio da super amostragem. Uma imagem de mais alta resolução pode ser reconstruída a partir de várias outras imagens de baixa resolução.

O poder de resolução de uma imagem pode ser aumentado por meio da inserção de informações de mais alta frequência tipicamente baseadas no conhecimento sobre o modelo específico de imagem (Vandewalle *et. al* – 2005).

O termo super-resolução não é muito recente e foi introduzido pela primeira vez por Tsai e Huang – 1984.

A maioria dos métodos para super-resolução é composto por duas etapas: Inicialmente todas as imagens de baixa resolução são alinhadas em um mesmo *grid* (*registration*) e em seguida uma imagem de mais alta resolução é construída a partir de uma relação matemática integrando todas ou parte das imagens de baixa resolução já alinhadas (reconstrução).

O método proposto por este trabalho está focado somente no processo de reconstrução.

Este artigo apresenta um modelo simples e eficiente para o aumento da resolução de uma imagem a partir de imagens já alinhadas. Os resultados são comparados com os métodos *Pocs* proposto por Papoulis e Gerchberg – 1977, *Iterated Back Projection* proposto por Irani e Peleg – 1991, *Robust Super Resolution* proposto por Zomet *et al.* – 2001, uma outra versão alterada do *Pocs* e *Structure-Adaptive Normalized Convolution* proposto por Pham *et al.* – 2006.

2. Estado da Arte

Papoulis e Gerchberg – 1977 propõem uma primeira versão do que viria a ser o algoritmo *Pocs* que basicamente projeta todos os pixels alinhados no mesmo *grid*, realiza uma filtragem no domínio da frequência e avalia o resultado. O filtro no domínio da frequência pode ser um passa faixas ou passa baixas. O resultado é avaliado para identificar uma possível convergência para a imagem de alta resolução. Caso o resultado da iteração anterior esteja muito próximo da iteração atual o método pára.

Já Irani e Peleg – 1991 propõem o algoritmo *Iterated Back Projection*. A idéia por trás desse algoritmo é basicamente iniciar com uma estimativa ruim da imagem de alta resolução e iterativamente adicionar uma nova imagem gradiente ao processo. Essa imagem gradiente seria

a soma de todos os erros entre a imagem de baixa resolução e a imagem de alta resolução estimada pelo processo de transformação dado pela estimativa de movimento entre as imagens de baixa resolução.

Zoomet *et al.* –2001 propõem o algoritmo *Robust Super Resolution* que é basicamente uma versão melhorada do *Iterated Back Projection* onde a única diferença reside no fato de que ao somar todos os erros para a estimação da imagem de alta resolução, esse método utiliza como imagem gradiente a média de todos os erros.

O algoritmo Pocs (*projection onto convex sets*) é uma variação do primeiro proposto por Papoulis e Gerchberg, entretanto o processo de filtragem no domínio da frequência para esse método é um filtro passa baixas projetado de tal forma para aproximar-se da PSF (*point spread function*) da câmera.

O algoritmo *Structure-Adaptive Normalized Convolution* proposto por Pham *et al.* – 2006 usa convolução normalizada para reconstruir a imagem de alta resolução. Duas opções podem ser habilitadas para esse algoritmo: Robustez a ruídos que basicamente analisa todas as imagens de baixa resolução e decide quais pixels podem estar ruidosos e não os considera no processamento e um segundo processamento para realizar a correção de possíveis erros. Essa segunda etapa basicamente adapta o tamanho e a orientação dos filtros gaussianos na convolução normalizada. Isso pode resultar em uma imagem mais realçada.

3. Modelo Proposto – composição de imagens realçadas

O modelo proposto insere informações de mais alta frequência obtidas a partir de vários quadros em um único quadro objetivando obter uma imagem com maior poder de resolução.

Suponha que existam p quadros de baixa resolução alinhados. Seja F um *kernel* para um filtro gaussiano. Seja I qualquer imagem de baixa resolução que faz parte do conjunto dos p quadros. Logo $I_{altas} = I - I * F$. Onde I_{altas} representa um mapa das altas frequências de uma imagem I do conjunto contendo os p quadros.

Pela própria natureza do processo de formação de imagens tanto em câmeras com sensores do tipo CCD quanto CMOS, é fato que dois quadros consecutivos da mesma cena apresentam informações de alta frequência diferentes.

Pode-se obter cada uma das imagens de baixa resolução com altas frequências realçadas realizando um processamento do tipo: $I_{realce} = I + I_{altas}$.

Em notação lexicográfica uma imagem I_k de dimensões $M \times N$ pode ser expressa da seguinte

maneira $I_k = [i_{k,1}, i_{k,2}, i_{k,3}, \dots, i_{k,Z}]^T$, onde $Z = M \times N$. Seja x o fator de ampliação da imagem e suponha que todas as imagens do conjunto p estejam com as altas frequências realçadas. Logo um quadro de mais alto poder de resolução pode ser composto da seguinte maneira: $I_{AR} = [i_{k,1}, i_{k+1,1}, i_{k+2,1}, \dots, i_{k+x,1}, i_{k,2}, i_{k+1,2}, i_{k+2,2}, \dots, i_{k+x,2}, i_{k,Z}, i_{k+1,Z}, i_{k+2,Z}, \dots, i_{k+x,Z}]^T$.

Esse processo pode ser repetido n vezes tomando subconjuntos de x em x imagens e ao final pode-se integrar todas as imagens de mais alta resolução em apenas uma.

4. Resultados

Os resultados serão mostrados e comparados com os resultados apresentados pelos principais métodos descritos acima.

Uma imagem de alta resolução foi tomada como referência para comparação com as reconstruções. Portanto todas as imagens de baixa resolução tomadas como entrada para todos os métodos são versões degradadas de imagens originais com mais alto poder de resolução. A **Figura 1** representa uma dessas imagens de alta resolução.



Figura 1 – Imagem de mais alta resolução

A **Figura 2** apresenta uma imagem de baixa resolução degradada a partir da **Figura 1**.



Figura 2 – Imagem de baixa resolução

As **Figuras 3 e 4** apresentam a representação das **Figuras 1 e 2** respectivamente no domínio da frequência.

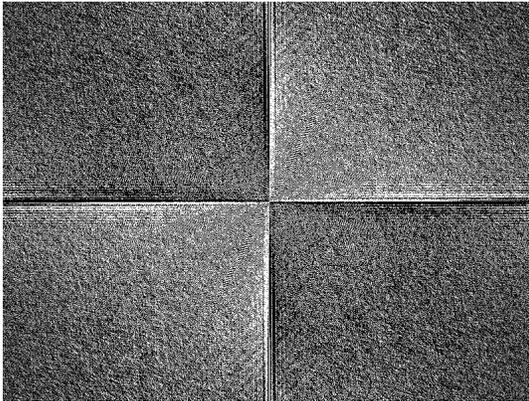


Figura 3 – FFT da imagem de alta resolução.

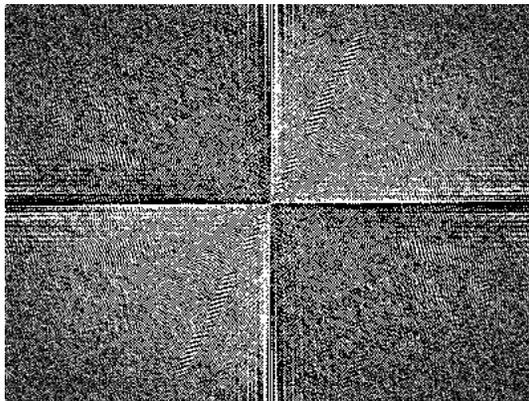


Figura 4 – FFT da imagem de baixa resolução.

É pertinente observar que a imagem de mais alta resolução possui uma maior quantidade de componentes de mais alta frequência como era de se esperar (resultado apresentado por Vandewalle *et. al* – 2005).

O resultado apresentado pelo método de Papoulis e Gerchberg – 1977 é mostrado na **Figura 5** e sua representação no domínio da frequência é a **Figura 6**.



Figura 5 – Resultado de Papoulis e Gerchberg

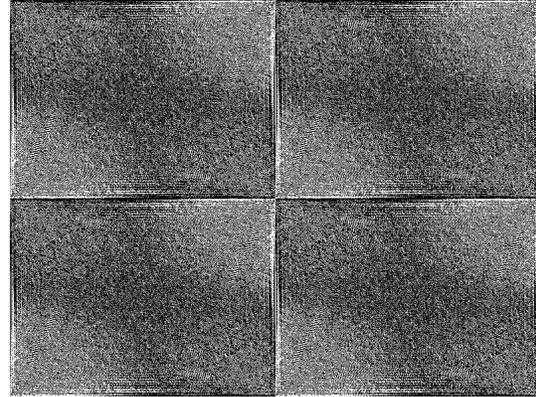


Figura 6 – FFT do resultado proposto por Papoulis e Gerchberg

Esse algoritmo aumentou as altas frequências, mostrando uma melhoria na imagem, entretanto percebe-se uma queda no nível geral de intensidade da imagem.

O resultado apresentado pelo método de Irani e Peleg – 1991 (*Iterated Back Projection*) é mostrado na **Figura 7** e sua representação no domínio da frequência é a **Figura 8**.



Figura 7 – Resultado de *Iterated Back Projection*

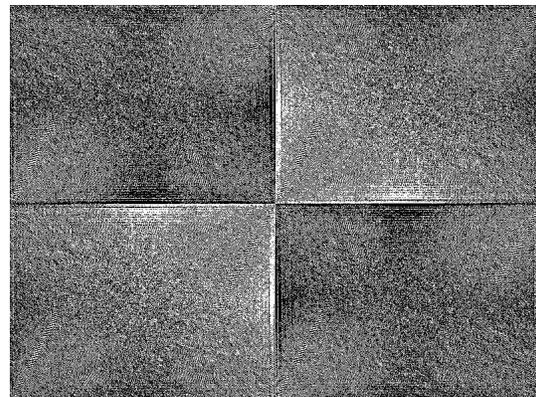


Figura 8 – FFT do resultado de *Iterated Back Projection*

Neste último resultado vale ressaltar que houve um aumento ainda maior das altas frequências e da qualidade da imagem reconstruída quando comparado com o resultado anterior de Papoulis e Gerchberg.

O resultado apresentado pelo método de Zoomet *et al.* –2001 (*Robust Super Resolution*) é mostrado na **Figura 9** e sua representação no domínio da frequência é a **Figura 10**.



Figura 9 – Resultado do método *Robust Super Resolution*

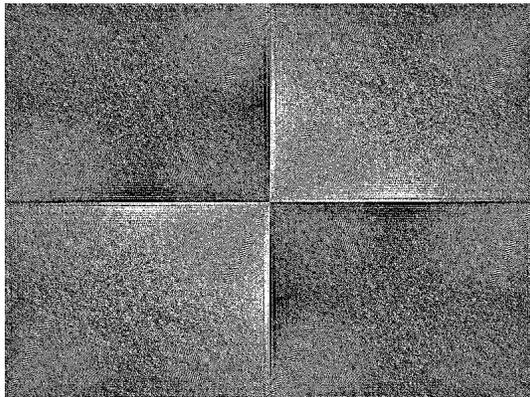


Figura 10 – FFT do resultado do método *Robust Super Resolution*

Os resultados apresentados por esse método são muito parecidos com os resultados apresentados pelo método anterior (*Iterated Back Projection*). Isso se deve à semelhança entre os dois algoritmos, sendo esse último uma evolução do primeiro. Apenas pela análise do espectro de frequência fica difícil identificar qual apresentou um melhor resultado.

O resultado apresentado pelo algoritmo Pocs melhorado (utilizando uma simulação da PSF) (*Iterated Back Projection*) é mostrado na **Figura 11** e sua representação no domínio da frequência é a **Figura 12**.



Figura 11 – Resultado do método *Pocs* com PSF

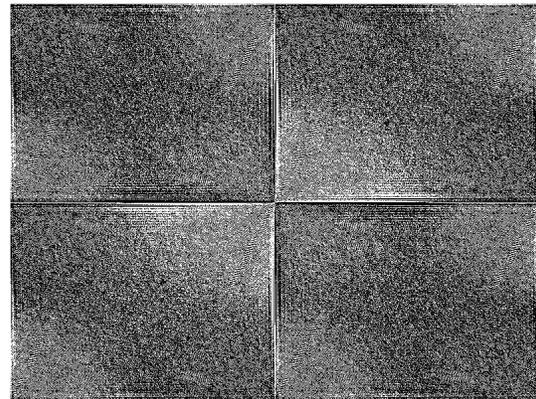


Figura 12 – FFT do resultado do método *Pocs* com PSF

Apesar desse método ser uma evolução do primeiro proposto por Papoulis e Gerchberg, ao repararmos os resultados entre os dois, no domínio da frequência, percebe-se um aumento na intensidade das altas frequências. Entretanto é ainda difícil quantizar com precisão a melhoria.

O resultado apresentado pelo método *Structure-Adaptive Normalized Convolution* proposto por Pham *et al.* – 2006 é mostrado na **Figura 13** e sua representação no domínio da frequência é a **Figura 14**.

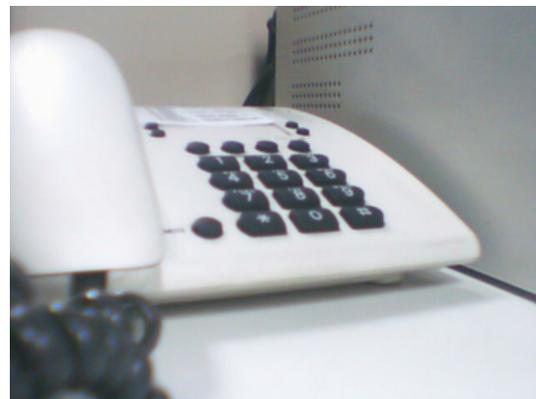


Figura 13 – Resultado do método *Structure-Adaptive Normalized Convolution*

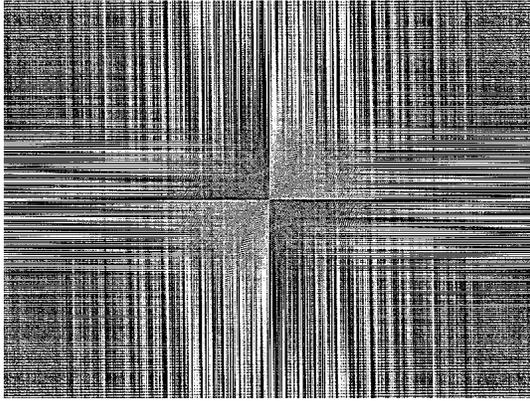


Figura 14 – FFT do resultado do método *Structure-Adaptive Normalized Convolution*

O processo de convolução normalizada inseriu padrões de alta frequência facilmente identificáveis no domínio FFT. O resultado na imagem original foi bom, entretanto esse foi o método que apresentou maior diferença entre o espectro de frequência original e da imagem reconstruída.

O resultado apresentado pelo algoritmo de aumento de resolução baseado em composição de imagens realçadas é mostrado na **Figura 15** e sua representação no domínio da frequência é a **Figura 16**.



Figura 15 – Resultado algoritmo composição por imagens realçadas.

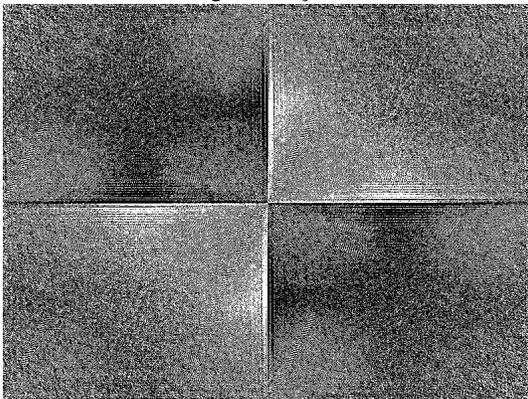


Figura 16 – FFT do resultado algoritmo composição por imagens realçadas.

Os resultados apresentados por esse último método mostram um aumento das componentes de alta frequência no domínio da frequência. Isso fornece subsídios de que o poder de resolução da imagem aumentou.

A próxima seção apresentará uma análise dos resultados com intuito de identificar uma métrica quantitativa de melhorias entre os algoritmos já apresentados.

5. Análise dos Resultados

Por definição, o espectro de potência fornece a quantidade de energia carregada por dada frequência. Portanto, um método de avaliação dos resultados poderia ser uma análise do espectro de potência de cada uma das imagens reconstruídas. Assim sendo, uma sugestão para a identificação da melhoria da resolução residiria no fato de que imagens com mais energia em frequências altas apresentam maior poder de resolução. Para tanto serão apresentados os dois espectros de potencia, tanto da imagem original de alta resolução quanto da imagem degradada de baixa resolução.

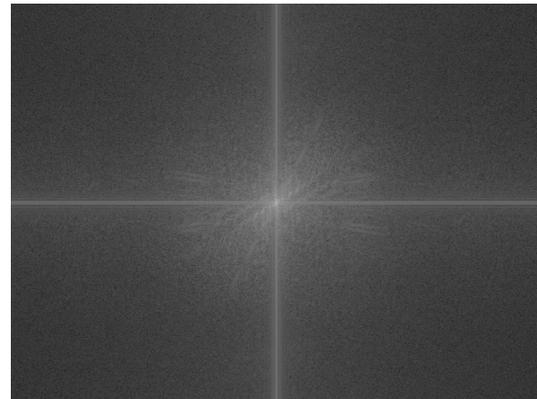


Figura 17 – Espectro de potência da imagem de alta resolução original.

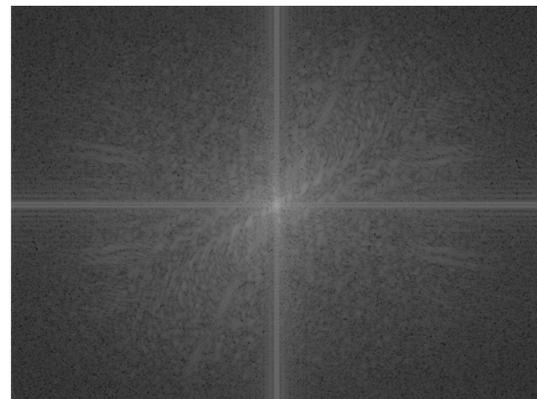


Figura 18 – Espectro de potência da imagem de alta resolução original.

É pertinente observar que a **Figura 18** mostra somente as frequências da região central da imagem original.

As próximas figuras mostram o resultado do espectro de frequência dos algoritmos *Pocs*, *Iterated Back Projection*, *Robust Super Resolution*, *Pocs (PSF)*, *Structure-Adaptive Normalized Convolution* e Composição por imagens realçadas respectivamente.

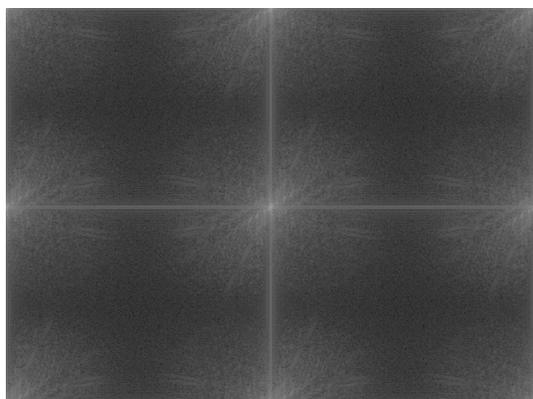


Figura 19 – Espectro de potência da imagem de alta resolução – *Pocs*.

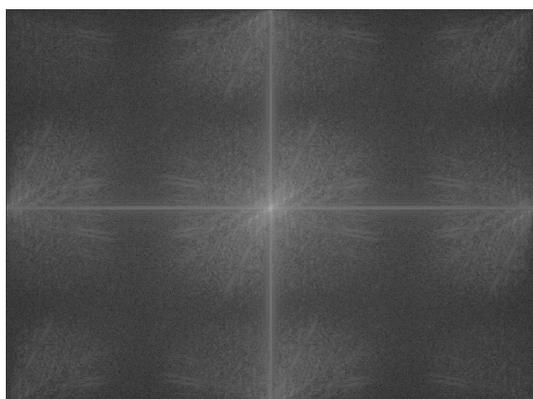


Figura 20 – Espectro de potência da imagem de alta resolução – *Iterated Back Projection*.

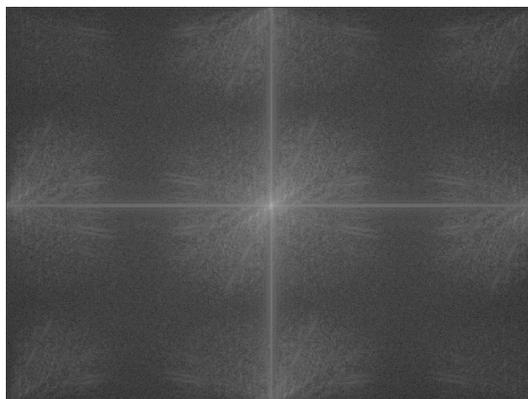


Figura 21 – Espectro de potência da imagem de alta resolução – *Iterated Back Projection*.

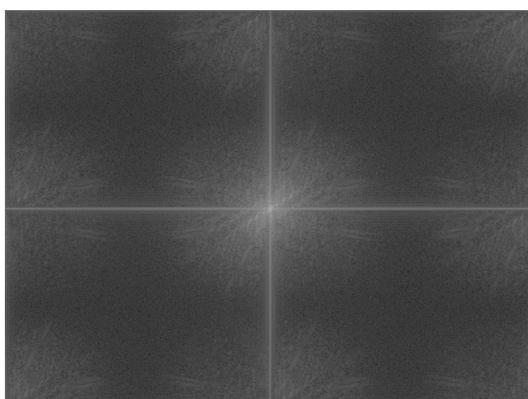


Figura 22 – Espectro de potência da imagem de alta resolução – *Pocs com PSF*.

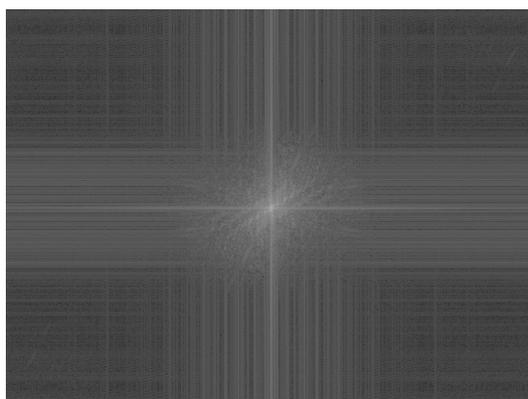


Figura 23 – Espectro de potência da imagem de alta resolução – *Structure-Adaptive Normalized Convolution*.

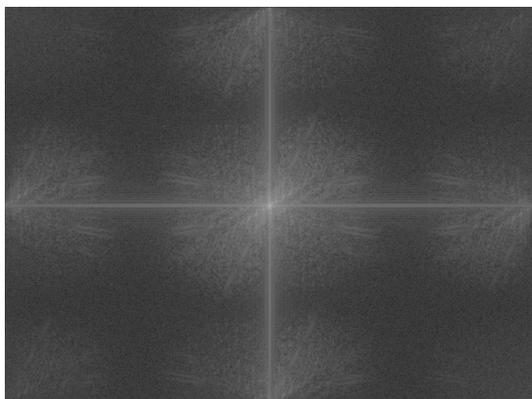


Figura 24 – Espectro de potência da imagem do algoritmo composição por imagens realçadas

É observável pelo espectro de potência que todos os métodos de super-resolução estudados apresentam energia em frequências mais altas que não existiam no espectro da imagem de baixa resolução.

Visualmente falando a partir do resultado da imagem reconstruída, o método *Structure-Adaptive Normalized Convolution* apresentou um resultado mais próximo da imagem de alta resolução original. Analisando o espectro de potência dessa abordagem, pode-se perceber uma energia maior sendo carregada por frequências mais elevadas que não estão presentes nos espectros dos demais métodos. O método proposto “composição por imagens realçadas”, apresentou-se melhor que o primeiro método proposto por Papoulis e Gerchberg.

O método *Iterated Back Projection* foi o que apresentou o segundo melhor resultado em termos de resolução depois do *Structure-Adaptive Normalized Convolution*.

Todavia, o método que apresentou um menor tempo de execução foi o proposto “composição por imagens realçadas” e o maior tempo de execução foi do algoritmo *Structure-Adaptive Normalized Convolution*.

6. Conclusões

É fato que é possível melhorar a resolução de uma imagem a partir de várias observações da mesma cena com imagens de baixa resolução.

O método *Structure-Adaptive Normalized Convolution* foi o que apresentou melhor resultado em termos de resolução da imagem reconstruída quando comparada com a imagem de alta resolução original.

O método proposto por Papoulis e Gerchberg foi o que apresentou o pior resultado em termos de resolução da imagem reconstruída quando comparada com a imagem de alta resolução original.

O método proposto “composição por imagens realçadas” pode contribuir para o aumento da resolução em aplicações em que o tempo de resposta é crítico.

7. Referências

[1] P. Vandewalle, S. Süssstrunk and M. Vetterli, A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution, EURASIP Journal on Applied Signal Processing (special issue on Super-resolution), Vol. 2006, pp. Article ID 71459, 14 pages, 2006.

[2] D. Keren, S. Peleg, and R. Brada, Image sequence enhancement using sub-pixel displacement, in Proceedings IEEE Conference on Computer Vision and Pattern Recognition, June 1988, pp. 742-746.

[3] Tuan Q. Pham, Lucas J. van Vliet and Klamer Schutte, Robust Fusion of Irregularly Sampled Data Using Adaptive Normalized Convolution, EURASIP Journal on Applied Signal Processing, Vol. 2006, Article ID 83268, 12 pages, 2006.

[4] M. Irani and S. Peleg, Improving resolution by image registration, Graphical Models and Image Processing, 53:231-239, 1991.

[5] A. Zomet, A. Rav-Acha, and S. Peleg, Robust Super-Resolution, Proceedings international conference on computer vision and pattern recognition (CVPR), 2001.

[6] A. Papoulis, Generalized sampling expansion, IEEE Transactions on Circuits Systems, vol. 24, no. 11, pp. 652-654, 1977.