

# Super-Resolução de Imagens de Sensores Remotos Usando a Transformada de Cor YUV para Estimação das Altas Frequências

Cléber Rubert  
Instituto Nacional de Pesquisas  
Espaciais - INPE  
cleber@dpi.inpe.br

Leila Maria Garcia Fonseca  
Instituto Nacional de Pesquisas  
Espaciais - INPE  
leila@dpi.inpe.br

Luiz Velho  
Instituto de Matemática Pura e  
Aplicada- IMPA  
lvelho@impa.br

## Resumo

*Este trabalho apresenta uma abordagem para a síntese de imagens com resolução espacial melhor do a imagem original, através do método chamado de super-resolução. O algoritmo baseia-se em amostras de treinamento para estimar detalhes em altas frequências nas imagens reamostradas. Neste trabalho o algoritmo é adaptado para usar a componente de luminância Y da transformada YUV para estimar as altas frequências perdidas no processo de reamostragem das imagens. Os resultados são comparáveis aos da abordagem que usa composição RGB, com a vantagem da redução da dimensionalidade do conjunto de amostras de treinamento e, conseqüentemente, do tempo computacional.*

**Palavras-chave:** *super-resolução, resolução espacial, modelo YUV; síntese de imagens, qualidade visual.*

## 1. Introdução

Em várias aplicações na área de sensoriamento remoto, o processo de reamostragem da imagem em uma grade mais fina é uma operação bastante utilizada. Neste processo (ampliação da imagem), pixels novos são criados, o que pode levar a uma degradação significativa na qualidade visual da mesma, apresentando, geralmente, uma aparência “borrada” devido à técnica de interpolação usada. Este borramento está relacionado com a perda de detalhes, os quais estão diretamente relacionados com o conteúdo em altas frequências da cena [14].

Para resolver esse problema, técnicas que utilizam abordagens estatísticas e aprendizagem de máquina (*machine learning*) [7] para a estimação das altas frequências têm sido propostas na literatura. Estas técnicas são chamadas de super-resolução. O objetivo destas técnicas é estimar a informação de detalhes na imagem reamostrada e que não podem se tornar visíveis através de métodos simples de realce. Esses algoritmos tentam capturar nuances de uma dada amostra, como

por exemplo, arestas, bordas, texturas, etc., para usar como informação para sintetizar uma imagem de melhor resolução.

O método proposto neste trabalho é uma adaptação da abordagem proposta por Freeman et. al. [6]. Esta metodologia tem sido adaptada para trabalhar com a componente de luminância da transformada YUV, com objetivo de reduzir o tempo computacional do processo de estimação das altas frequências. A estimação das altas frequências é realizada usando imagens de diferentes sensores com melhor resolução espacial do que a da imagem de entrada.

### 1.1 Super-resolução

Super-resolução está relacionada com técnicas de extrapolação da resolução espacial das imagens digitais e, por conseqüente, na melhoria da qualidade visual das mesmas. Basicamente, há três formas de melhorar a resolução aparente de uma imagem:

1) Realce de detalhes existentes na imagem. Esta operação é realizada através da amplificação das componentes em altas-frequências já presentes no sinal, através de técnicas como restauração [2] e/ou como em técnicas de interpolação adaptativa [1 e 4]. Nestes casos são utilizados filtros adaptativos que amplificam as componentes de alta frequência atenuadas no processo da reamostragem; em alguns casos, aplica-se posteriormente um pós-processamento para reduzir os artefatos inseridos nas imagens interpoladas.

2) Síntese de uma única imagem de alta resolução a partir de uma seqüência de imagens de baixa resolução. É uma das técnicas mais estudadas em super-resolução. Alguns trabalhos abordam esta questão como um problema de reconstrução de imagens [3] [5] ou também como fusão de dados [12]. Esta abordagem permite recuperar detalhes reais (não extrapolados) das seqüências de imagens e adicioná-los na imagem estimada. De acordo com Park et. al. [11], os algoritmos de super-resolução a partir de múltiplos quadros, geralmente, passam por três fases distintas: i) registro das imagens de baixa resolução ou estimação do

deslocamento de cada imagem em relação a outra; ii) criação de uma grade de alta resolução, onde os valores da mesma são interpolados a partir das imagens registradas; iii) remoção de ruídos dependendo do modelo de observação.

3) Super-resolução baseado em aprendizagem. O objetivo é estimar a informação de detalhes que não está presente na imagem original através de amostras de treinamento. Em Freeman et. al. [6] e Hertzmann et. al. [10], são propostos algoritmos que usam algumas imagens de referências com o objetivo de treinar o algoritmo e assim sintetizar uma imagem com melhor qualidade visual do que a imagem original. Para sintetizar a imagem de super-resolução, o algoritmo extrai detalhes que não existem na imagem de baixa resolução, a partir de um conjunto de imagens de treinamento.

Embora as formas de melhorar a resolução espacial, apresentadas nos itens 1 e 2, sejam conhecidas também como super-resolução, neste trabalho o problema de super-resolução se refere ao problema de estimação da informação de detalhes na imagem ampliada, cuja abordagem se enquadra no terceiro caso.

Os passos deste método são ilustrados na Figura 1. Nesta abordagem, interpola-se a imagem de entrada para obter a imagem ampliada. No processo de ampliação as informações de altas frequências são perdidas ou atenuadas. Em um segundo passo, as informações de detalhes são estimadas usando um conjunto de amostras de treinamento. Esses detalhes estimados são adicionados à imagem de entrada interpolada. O algoritmo “aprende” a relação existente entre os blocos de baixas e altas frequências e assim, estima as informações de alta frequência na imagem de baixa frequência.

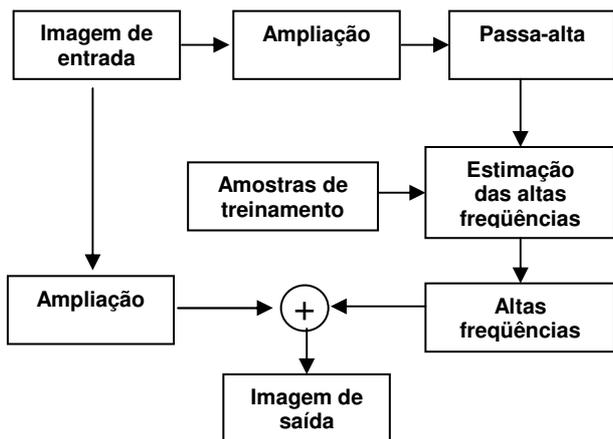


Figura 1. Passos do algoritmo de super-resolução.

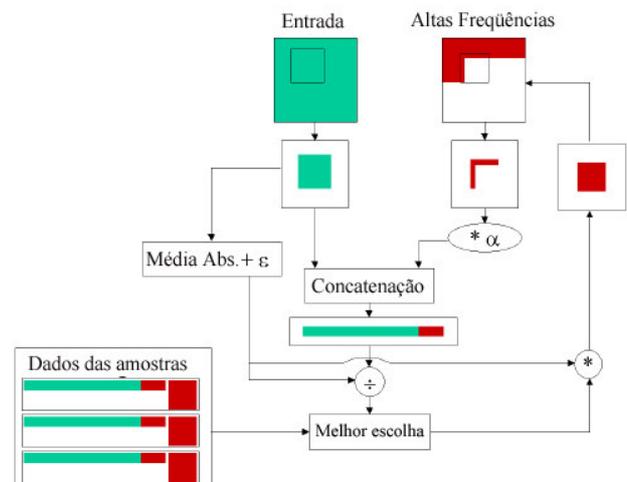
O conjunto de treinamento é composto por pares de bandas de médias e altas frequências. Assim, versões de imagens de baixa (degradada), média e alta resoluções são obtidas através da decomposição em multi-resolução.

A imagem de baixa frequência é obtida aplicando

uma filtragem passa-baixa, reduzindo e depois ampliando, onde temos então a versão degradada da imagem original. A banda de alta frequência é gerada através da operação de diferença entre a imagem original e a sua versão degradada, enquanto a banda de média frequência é extraída da imagem degradada aplicando um filtro passa-alta.

A seguir, as imagens são particionadas em blocos. O conjunto de treinamento é composto pelos blocos de média e alta resolução. Esses blocos, não precisam ser necessariamente do mesmo tamanho. Neste trabalho, são usados blocos de 7x7 e 5x5, para as imagens de médias e altas frequências, respectivamente. A escolha do tamanho do bloco é um compromisso entre dois extremos. Se o tamanho do bloco da imagem for muito pequeno, o local deste na imagem dará muita pouca informação para estimar as frequências da cena. Por outro lado, se o tamanho do bloco for muito grande, a base de dados necessária para estimar as altas frequências cresce exponencialmente com a dimensionalidade dos blocos de imagens [6].

A Figura 2 apresenta a seqüência do algoritmo para estimar a informação de altas frequências na imagem de super-resolução. A estimação do bloco de alta frequência (vermelho) é uma fase muito importante do algoritmo, onde dois problemas devem ser avaliados. O primeiro deles é o problema de busca que consiste em encontrar, no conjunto de treinamento, o bloco de alta resolução que seja mais similar ao bloco analisado da imagem de entrada (verde). Outro problema é como preservar a continuidade da informação nas bordas do bloco que está sendo analisado. Para isso usa-se uma sobreposição da imagem já estimada. Esta sobreposição é multiplicada por um parâmetro  $\alpha$ , para garantir que o bloco selecionado no conjunto de treinamento seja coerente com os blocos vizinhos. Os valores do bloco de baixa resolução e da sobreposição são concatenados no vetor de busca.



Devido à complexidade em representar a diversidade e a riqueza de informação de uma imagem natural, supõe-se que o relacionamento estatístico entre as bandas é independente do contraste. Para isso, os dados do conjunto de treinamento são gerados a partir das imagens normalizadas pelo valor absoluto da banda de médias frequências somados ao pequeno valor  $\mathcal{E}$  para evitar divisão por zero.

## 1.2 Modelo YUV de cores

O modelo YUV é usado na transmissão comercial de TV colorida, onde o modelo RGB é recodificado para a eficiência de transmissão e para manutenção de compatibilidade como os padrões monocromáticos de TV. Este modelo contém uma componente de luminância (Y) e duas componentes de cores (U & V). A componente Y contém a informação de luminância (brilho = variações em tons de cinza) e as componentes U e V representam a informação de cor [8]. A conversão do sistema RGB para YUV é dada pela seguinte equação:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

A principal vantagem do modelo YUV em processamento de imagem é que a luminância e a informação de cor são independentes. Assim, a componente de luminância pode ser processada sem afetar seu conteúdo de cor [8].

Acredita-se que as informações de detalhes, em uma imagem digital, estejam presentes na luminância da cena. Por isso, neste trabalho, utilizou-se a componente Y para estimação das altas frequências, e com isso, diminuir a dimensionalidade do vetor de busca e do conjunto de treinamento.

Assim, temos a vantagem da alta sensibilidade do sistema visual humano à variação de luminosidade do que da variação dos valores de crominância [8]. Conseqüentemente, mais recursos computacionais são alocados para ampliar os valores de brilho, enquanto as componentes de cores podem ser ampliadas usando uma abordagem mais simples.

## 2. Método

O método proposto foi implementado e avaliado usando imagens do satélite Landsat e Ikonos. Foram escolhidas imagens que contêm áreas urbanas, que possuem uma grande quantidade de detalhes, permitindo, assim, testar a eficiência do algoritmo em termos de preservação ou recuperação da informação de detalhes.

As componentes de luminância (Y) das imagens usadas no experimento (teste) foram decompostas

utilizando a pirâmide laplaciana [13]. O conjunto das amostras de treinamento é formado, desta forma, pelas bandas de médias e altas frequências obtidas no primeiro nível da decomposição.

No algoritmo de Freeman et. al. [6], os valores do vetor de busca são definidos a partir do três canais RGB dos blocos da imagem. Neste trabalho é proposto o uso da informação de luminância da cena (sistema de cores YUV) para criar o vetor de busca na base de dados de amostras. A idéia pe reduzir o custo computacional do processo de busca.

Com a abordagem proposta, o tamanho do vetor de busca que continha 174 posições, no caso da composição RGB foi reduzido para 58 posições usando somente a componente Y do modelo YUV. Assim, o tamanho do espaço de armazenamento do conjunto de treinamento foi também reduzido.

As frequências estimadas são adicionadas à componente Y da imagem de entrada ampliada. As componentes de cor U e V são ampliadas usando uma interpolação bilinear. Depois de estimadas as frequências o modelo é convertido novamente para o espaço RGB para gerar a imagem resultante.

## 3. Resultados

A Figura 3 apresenta os resultados do uso do algoritmo de super-resolução para uma imagem Ikonos, utilizando como treinamento as imagens mostradas na Figura 3(d). Como podemos observar, a imagem da Figura 3(c) possui uma aparência mais nítida que a imagem interpolada mostrada na Figura 3(b). Das técnicas de interpolação tradicionais, o interpolador bicúbico é que mais preserva os detalhes da imagem, e por isso foi utilizado como um parâmetro de comparação visual com o método de super-resolução.

A Figura 4 mostra o resultado da aplicação do método de super-resolução aplicado a uma imagem Landsat para fins de análise do método proposto. O processamento foi realizado nos seguintes passos: o tamanho da imagem original é reduzida por um fator de 2; a imagem sub-amostrada é reamostrada para seu tamanho original usando o interpolador bicúbico e pelos métodos de super-resolução, para fim de comparação. Das Figuras 4(b) e 4(c), pode-se observar que a imagem de super-resolução apresenta a informação de detalhes mais nítida do que na imagem interpolada mostrada na Figura 4(c). As diferenças entre a imagem interpolada e a imagem de super-resolução se manifestam com maior nitidez nas bordas e nas regiões de textura da imagem.

A Tabela 1 apresenta as medidas de qualidade de imagens, dado pelo Erro Médio Quadrático (MSE) e pelo Índice de Qualidade de Imagem Universal (IQI) proposta por [15].

**Tabela 1. Avaliação quantitativa dos métodos de Super-resolução.**

<i>Método</i>	<i>MSE</i>	<i>IQI</i>
SR (RGB)	460,47	0,9765
SR (YUV)	418,12	0,9772

Pode-se observar que o método de super-resolução proposto, usando a componente Y, ficou próximo do método utilizando a composição RGB, já que para este caso, o IQI ficou mais próximo a 1 (um) e o MSE ficou menor em relação ao método com RGB.

O perfil radiométrico da imagem também foi usado para analisar a informação de detalhes inseridos no processamento. A Figura 5 mostra o perfil radiométrico de uma linha das imagens processadas pelo interpolador bicúbico e pelos métodos de super-resolução. As variações abruptas na imagem estimada com super-resolução vistas nos gráficos da Figura 5 (c) e (d) podem ser relacionadas com detalhes na imagem. O método de super-resolução teve um resultado mais próximo do que o interpolar bicúbico (b) em relação à imagem original (a). Por outro lado, pode ser observado, em algumas regiões, a presença de alguns artefatos inseridos no processamento.

A Figura 6, mostra o espectro de Fourier das imagens processadas mostradas na Figura 4. Pode-se observar que os métodos de super-resolução amplificaram as componentes de altas frequências de acordo com se espectro mostrado na Figura 6(c) e (d), ficando mais semelhante ao da imagem original (a), comparado com o espectro da imagem interpolada pelo método bicúbico apresentado na Figura 6(b), onde as altas frequências possuem uma magnitude menor devido à atenuação das componentes de alta frequência no processo de interpolação. Como mostra o espectro, as imagens estimadas com super-resolução possuem mais detalhes que a imagem interpolada.

#### 4. Conclusões

Este trabalho apresentou uma proposta de usar a componente Y do modelo de cores YUV, ao invés das três componentes do sistema de cores RGB na síntese da imagem de super-resolução. As imagens processadas pelo método proposto apresentaram resultados similares ao método que utiliza a composição RGB. Além disso, o método apresenta a vantagem de ser muito mais barato computacionalmente, uma vez que reduz a dimensionalidade do conjunto de treinamento e do vetor de busca.

Com este propósito, em trabalhos futuros pretende-se usar bandas pancromáticas, que possuem alta resolução como imagens de treinamento e assim gerar uma imagem colorida de melhor resolução, a partir de um conjunto de imagens multi-espectrais de baixa-resolução.

Em termos computacionais, geralmente, as abordagens de super-resolução são caras. O número de

blocos no conjunto de treinamento e as dimensões dos vetores de busca contribuem para que o processo de busca fique lento computacionalmente. Assim, o próximo passo dessa abordagem é implementar um algoritmo de busca que possa realizar o processamento de forma mais rápida, como, por exemplo, usar uma quantização vetorial (TSVQ), clusterização (ANN, SOM, etc).

#### Referências Bibliográficas

- [1] Battiato, S.; Gallo, G.; Stanco, F. "A locally adaptive zooming algorithm for digital images". *Image and Vision Computing*, v. 1, n. 20, p. 805-812, 2002.
- [2] Boggione, G. A.; Fonseca, L.M.G. "Restoration of Landsat-7 Images". In: *International Symposium on Remote Sensing of Environmental (ISRSE)*, Hawai, Nov. 2003.
- [3] Borman, S.; Stevenson, R. L. "Super-Resolution from Image Sequences - A Review". In: *Midwest Symposium on Circuits and Systems*, Notre Dame, IN. 1998.
- [4] Candocia, F. M.; Principe, J. C. "Super-resolution of images based on local correlations". *IEEE Transaction on Neural Networks*, v. 10, n. 2, p. 372-380, Mar 1999.
- [5] Capel, D.; Zisserman, "A. Super-resolution from multiple views using learnt image models". *IEEE - Computer Vision and Pattern Recognition*, v. 2, p. 627-634, 2001.
- [6] Freeman, W. T.; Jones, T. R.; Pasztor, E. C. "Example-based super-resolution". *IEEE Computer Graphics and Applications*, v. 1, n. 02, p. 56-65, March/April 2002.
- [7] Freeman, W. T.; Pasztor, E. C.; Carmichael, O. T. *Learning Low-level Vision*. Cambridge, MA: MERL, Mitsubishi Electric Research Labs, 2000. 45 p. (TR-2000-05a).
- [8] Gonzales, R.C. and Woods, R.E. *Digital Image Processing*. Addison-Wesley, 1993.
- [9] Bergen, J. R. "Pyramid-based Texture Analysis/Synthesis". In: *Proc. ACM SIGGRAPH*, Los Angeles, August, 1995.
- [10] Hertzmann, A.; Jacobs, C. E.; Oliver, N.; Curless, B.; Salesin, D. H. "Image Analogies". In: *SIGGRAPH*, 12-17 August, Los Angeles. *Proceedings of Conference on Computer Graphics and Interactive Techniques*. 2001.
- [11] Park, S. C.; Park, M. K.; Kang, M. G. "Super-Resolution Image Reconstruction: A Technical Overview". *IEEE Signal Processing Magazine*. n. 03. p. 21-36. 2003.
- [12] Rajan, D.; Chaudhuri, S. "Data fusion techniques for super-resolution imaging". *Information Fusion*, v. 3, p. 25-38, 2002.
- [13] Rosenfeld, A. *Multiresolution image processing and analysis*. Ed. Springer-Verlang, New York, 1984.
- [14] Sakamoto, T.; Nakanishi, C.; Hase, T. "Software pixel interpolation for digital still cameras suitable for a 32-bit MCU". *IEEE Transaction on Consumer Electronics*, v. 44, n. 4, p. 1342-1352, Nov 1998.
- [15] Wang, Z.; Bovik, A. C. "A universal image quality index". *IEEE Signal Processing Letters*, v. 9, n. 3, p. 81-84, March 2002.



Figura 3. Imagem Ikonos (a) re-amostrada com fator 2 a partir de uma interpolação bi-cúbica (b) e estimada com super-resolução (c) usando amostras de treinamento (d).

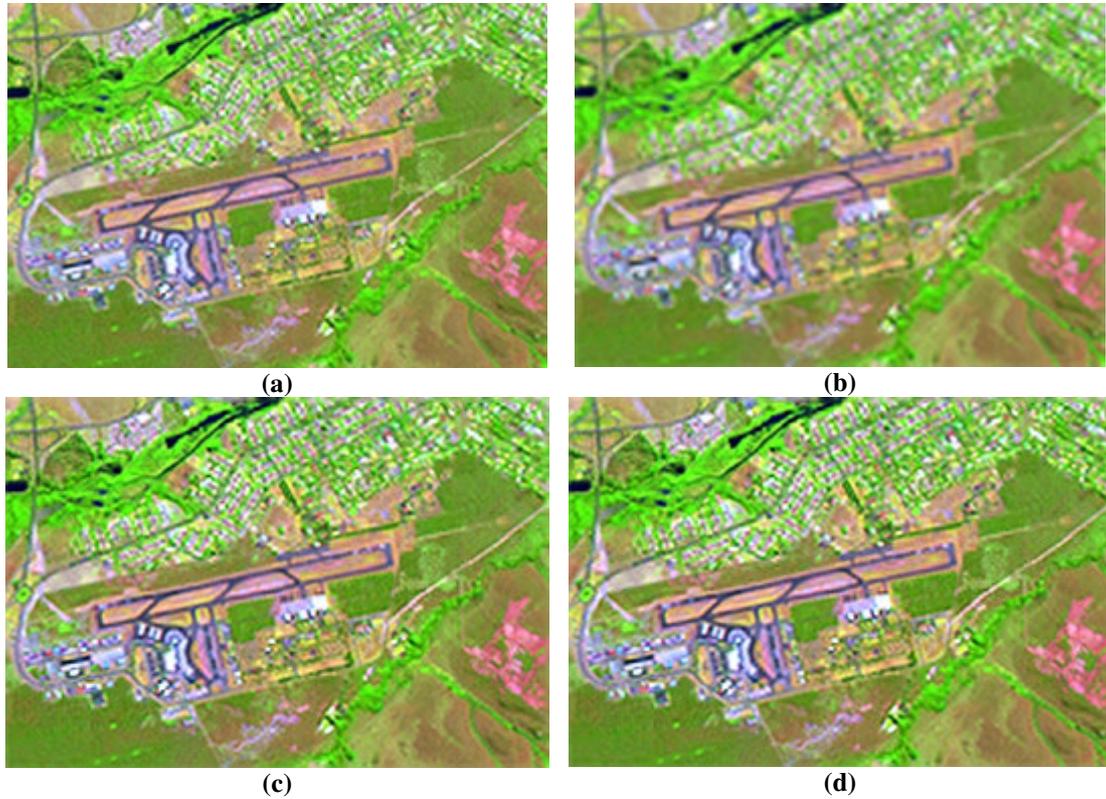


Figura 4. Imagem Landsat (a); re-amostrada com fator 2 a partir de um interpolador bicúbico (b), e estimada com super-resolução usando a componente YUV (c) e com super-resolução utilizando a composição RGB (d).

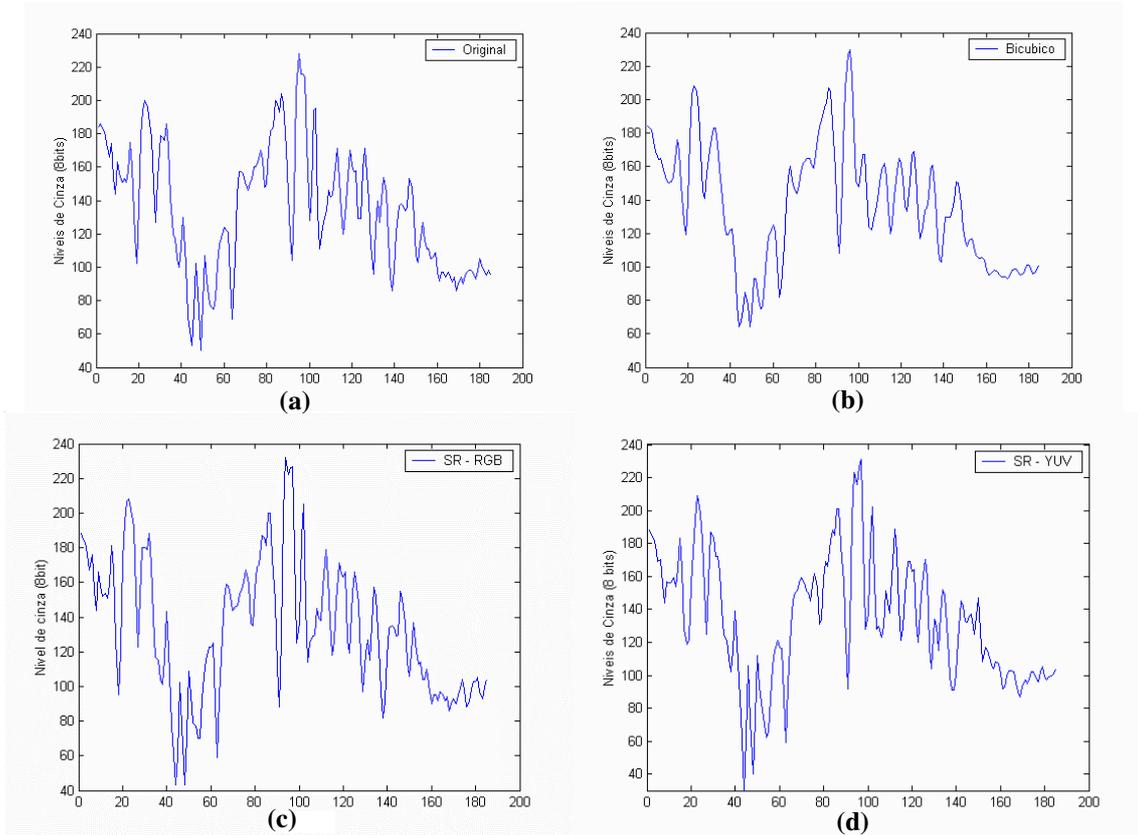


Figura 5. Perfil radiométrico da imagem Landsat ampliada com interpolador bicúbico (b) e com super-resolução pela composição RGB (c) e pela transformada YUV (d), comparado com o perfil da imagem original (a).

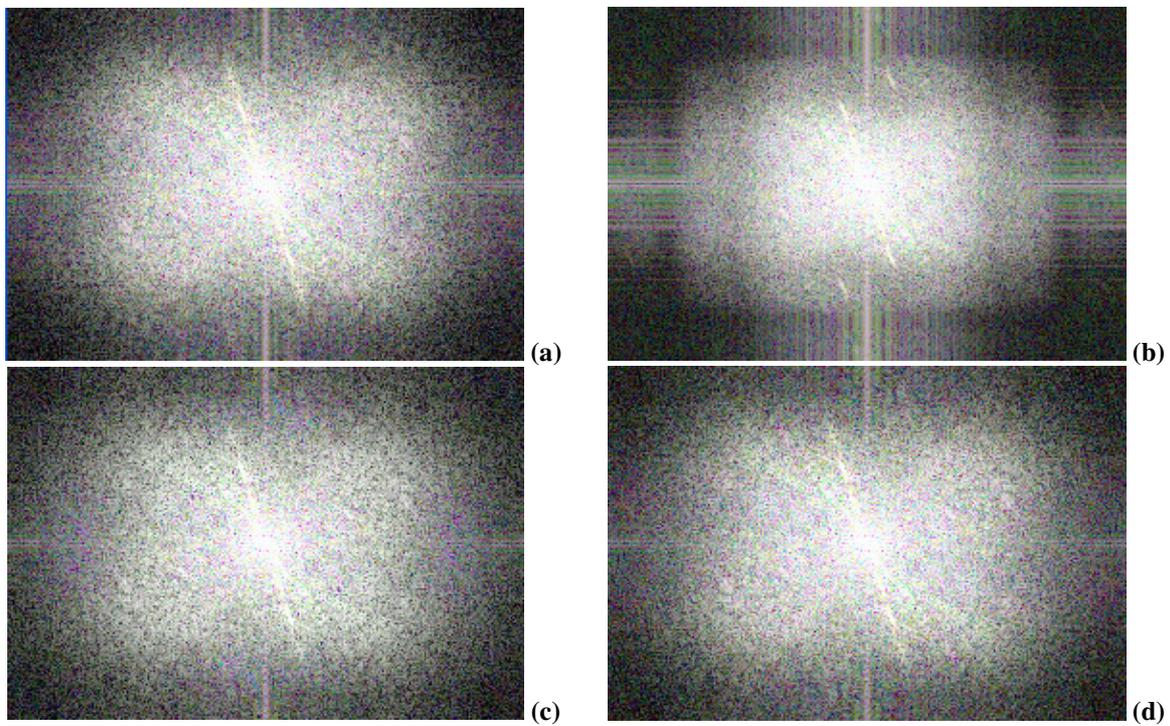


Figura 6. Espectro de Fourier para a imagem Landsat (a); ampliada com interpolador bicúbico (b); super-resolução pela transformada YUV (d) e com super-resolução pela composição RGB (c).