

DETECÇÃO DE LINHA DO HORIZONTE EM IMAGENS DIGITAIS

HORIZON LINE DETECTION IN DIGITAL IMAGES

BEUREN, Arlete Teresinha¹; FACON, Jacques²

¹Universidade Tecnológica Federal do Paraná – UTFPR, Santa Helena, Paraná, Brasil; ²Pontifícia Universidade Católica do Paraná – PUC, Curitiba, Paraná, Brasil
arlete_pr@yahoo.com.br; facon@ppgia.pucpr.br

Resumo

Uma abordagem de detecção automática de linha do horizonte em imagens coloridas é proposta neste artigo. A imagem original é inicialmente segmentada em duas classes no espaço de cor RGB, usando o classificador não supervisionado KMédio. A imagem resultado é em seguida analisada para determinar a linha do horizonte. Os experimentos realizados a partir de uma base de imagens coloridas mostraram as potencialidades da abordagem no que diz respeito à detecção automática de linha do horizonte.

Palavras-chave: Segmentação, Imagens Coloridas, Classificador não supervisionado, Binarização.

Abstract

We describe a segmentation approach to detect the horizon line in digital images. The original images are segmented into two classes by KMeans classifier with RGB color space. No previous process of correction or contrast and brightness filtering is employed. The experiments performed from 44 image database have showed the potentialities of the approach with regard to the automatic detection of horizon line in color images.

Key-words: Color images; Segmentation; KMeans classifier; Threshold.

1. Introdução

No caso de imagens digitais e vídeos, a linha do horizonte representa uma informação valiosa e é usada para diferentes fins, tais como auxiliar a navegação de veículos marinhos e aéreos, ajustar planos de vôos, e facilitar a vigilância para evitar acidentes, principalmente colisões com obstáculos. A detecção precisa da linha do horizonte em imagens contendo diversas informações como céu, prédios, estradas, gramado, mar, praias, continua sendo um desafio atual. A mistura de cores e padrões pode ser um empecilho na hora da detecção. A Figura 1 exemplifica uma imagem com a linha do horizonte destacada.

Figura 1 – Exemplo de detecção de linha do horizonte



Uma abordagem de detecção automática de linha do horizonte em imagens digitais é proposta neste artigo. As imagens originais sofrem nenhuma correção prévia de contraste e brilho e são segmentadas pelo classificador não supervisionado *KMédio* no espaço de cor *RGB*. Uma avaliação numérica da segmentação obtida é realizada para mensurar a eficiência da abordagem proposta.

2. Abordagem proposta

A detecção da linha do horizonte inicia-se segmentando as imagens em duas classes pelo classificador não supervisionado *KMédio* que agrupa indivíduos em função da similaridade entre eles, cada indivíduo sendo descrito por características.

1. Segmentação pelo *KMédio*: O algoritmo *KMédio* é um mecanismo de agrupamento que consiste em colocar os indivíduos estudados em grupos baseado na similaridade entre eles, cada indivíduo sendo descrito por características (LIAO, 2005). O algoritmo é baseado na minimização do índice de eficiência que é definido como a soma das distâncias quadradas de todos os pontos do cluster ao seu centro. Este procedimento consiste dos seguintes passos:

Passo 1: Escolher k centros iniciais das classes $z_1(1), z_2(1), \dots, z_k$;

Passo 2: Distribuir, no t -ésimo passo, as amostras $\{x\}$ entre as k classes, usando a relação da Equação 1:

$$X \in S_j(t) \text{ se } \|x - z_j(t)\| < \|x - z_i(t)\| \quad (1)$$

sendo $i = 1, 2, \dots, k, i \neq j$, onde $S_j(t)$ denota o conjunto de amostras cujo centro da classe é $z_j(t)$, no t -ésimo passo.

Passo 3: Calcular, a partir do resultado do passo 2, o novo centro da classe $z_j(t+1)$, $j = 1, 2, \dots, k$, tal que a soma das distâncias quadradas de todos os pontos em $S_j(k)$ para o novo centro do cluster seja minimizada, usando a relação que deve ser minimizada pela Equação 2:

$$J_j = \sum_{x \in S_j(k)} \|x - z_j(t+1)\|^2, j = 1, 2, \dots, k \quad (2)$$

sendo $z_j(t+1)$ o novo centro da classe. Consequentemente, o novo centro do cluster é dado pela Equação 3:

$$z_j(t+1) = 1/N_j \sum_{x \in S_j(t)} x, j = 1, 2, \dots, k \quad (3)$$

Passo 4: Verificar se $z_j(t+1) = z_j(t)$ para j

= 1, 2, ..., k. Neste caso o algoritmo convergiu e o procedimento é finalizado.

2. Características: Entre as várias características possíveis foram usadas no espaço de cor *RGB*:

- Cor: A cor em cada canal do *pixel* x_i é a primeira característica, conforme a Equação 4.

$$x_i = R \text{ ou } G \text{ ou } B \text{ e } x_i \in [0, 1, 2, \dots, 255] \text{ no domínio discreto (4)}$$

- Média: Permite caracterizar o comportamento local e reduzir a influência do ruído. Nos casos de homogeneidade absoluta da vizinhança, a média é igual ao valor do pixel, sendo N o número de *pixels* da região. A Equação 5 apresenta o cálculo da média μ para cada canal R , G e B .

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (5)$$

- Variância: Fornece a noção da distribuição dos níveis de cinza ao redor do pixel. Permite a caracterização do aspecto homogêneo, junto com a detecção da influência do ruído.

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

- *Skewness*: Definido como o quociente do momento de terceira ordem e a terceira potência do desvio padrão.

$$\gamma = \frac{1}{N\sigma^3} \sum_{i=1}^N (x_i - \mu)^3$$

- Curtoses: Definido como o quociente do momento de quarta ordem em torno da

média μ e a quarta potência do desvio padrão (σ).

$$c = \frac{1}{N\sigma^4} \sum_{i=1}^N (x_i - \mu)^4$$

3. Para detectar a linha do horizonte, as imagens foram previamente segmentadas usando o *KMédio* em duas classes, o céu e o chão, separadamente para cada canal R , G e B . Usou-se a distância euclidiana para classificar cada amostra no t -ésimo passo em função do centroide da classe (céu, não-céu) usando como características cada canal do *pixel* x_i e média μ correspondente (conforme Equações 4 e 5).

4. Detecção da linha do horizonte: cada imagem resultante da segmentação pelo *KMédio* é binarizada pelo algoritmo de Otsu (OTSU, 1979). Em seguida, pós processando as imagens empregando ferramentas morfológicas, a linha do horizonte é obtida por detecção de borda entre céu e chão.

3. Experimentos e Análise

Para testar a abordagem proposta, uma base de imagens coloridas foi criada a partir de uma base disponível *imageparsing*¹ contendo imagens coloridas de cidade, natureza, porto, rodovia, praia e paisagem rural. Interessou-se nesta base por também disponibilizar a solução ideal *groundtruth* para cada imagem da rotulação dos diversos itens semânticos. Desta base, uma nova base de 44 imagens foi gerada, cada imagem ideal *groundtruth* sendo reprocessada de maneira a gerar a imagem binária *groundtruth* da segmentação entre céu

¹ IMAGEPARSING, Disponível em:

http://www.imageparsing.com/Dataset_Album/, acesso em 30/07/2015.

e chão. Enquanto que a Figura apresenta exemplos de imagens *da base*, as Figuras 3(b),

4(b) e 5(b) ilustram exemplos de imagens com a solução *groundtruth*.

Figura 2 – Exemplos de imagens da base



Experimentos foram realizados utilizando-se a base gerada de 44 imagens com cada canal R, G e B separadamente. De maneira a avaliar o grau de eficiência de cada canal e de cada característica, experimentos foram realizados com cada canal R, G e B separadamente iniciando com duas características: cor e média.

Uma análise visual dos experimentos revelou que as grandes variabilidades de conteúdos resultaram em múltiplas segmentações céu-chão, algumas mais próximas da separação *groundtruth* céu-chão dependendo dos canais e dos conteúdos, dificultando um parecer final sobre qual ou quais canais eram mais interessantes. Percebeu-se então que somente a análise visual não podia ser conclusiva.

Decidiu-se, portanto, avaliar numericamente a eficiência da abordagem proposta comparando cada imagem binária final obtida com a imagem ideal correspondente (*groundtruth*). Para isto, métricas de especificamente elaboradas para avaliar a eficiência de técnicas de segmentação foram usadas: o erro de má classificação e o erro relativo da área do primeiro plano *RAE* (*Relative Area Error*) (SANKUR, 2001), as taxas de verdadeiros positivos e verdadeiros negativos $V+$ e $V-$, as taxas de falsos positivos e falsos negativos $F+$

e $F-$.

Também foram usadas a precisão P , a acurácia A , a revocação R e o erro ER (BIMBO, 1999). E por fim as métricas F -Measure FM e de taxa negativa NRM (*Negative Relative Metric*) (GATOS, 2009).

Os resultados da avaliação numérica da abordagem proposta para a base de 44 imagens estão reunidos na Tabela 1. Fica possível constatar que as taxas de erro de segmentação ME , RAE , FP , FN , ER e NRM são mais elevadas para os 2 canais R e G. E que as taxas de segmentação bem sucedida TP , TN , P , A , R e FM são mais elevadas para o canal B.

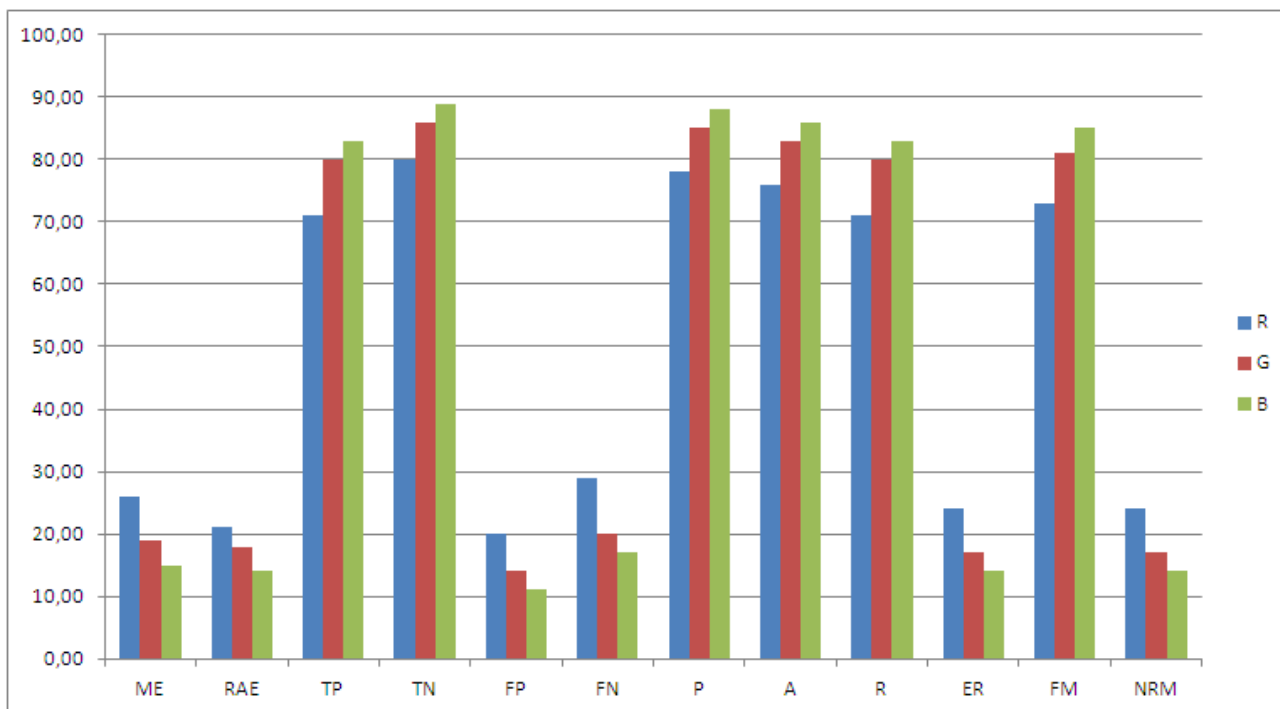
Os resultados da Tabela 1 evidenciaram ainda que o canal R é o menos interessante, com as maiores taxas de erro e as menores taxas de segmentação sucedida.

É possível ainda perceber que o canal B é o mais interessante dos 3 canais, com as menores taxas de erro e as maiores taxas de segmentação bem sucedida. Essa eficiência se reflete em todos os valores das métricas ME , RAE , TP , TN , FP , FN , P , A , R , ER . As taxas das métricas FM e NRM consideradas por [GATOS-2009] como sendo as mais criteriosas em termos de eficiência, são de 84.83% e 13.8% respectivamente. A Figura 3 ilustra os valores das métricas.

Tabela 1: Avaliação numérica da abordagem com as características, cor e média (em %).

	ME	RAE	TP	TN	FP	FN	P	A	R	ER	FM	NRM
Canal R	25.85	21.24	71.41	79.90	20.10	28.59	78.19	75.66	71.41	24.34	73.42	24.34
Canal G	19.17	17.73	79.56	85.92	14.08	20.44	84.69	82.74	79.56	17.26	81.21	17.26
Canal B	14.63	13.93	83.40	89.00	11.00	16.60	88.20	86.20	83.40	13.80	84.83	13.80

Figura 3: Visualização gráfica dos resultados da Tabela 2



A partir dos resultados numéricos descritos na Tabela 1, depois de ter definido o canal mais interessante, o canal B, foram acrescentadas gradativamente mais características cor + média + variância (C+M+V), cor + média + *Skewness* (C+M+VS) e cor + média + curtoses (C+M+C). Os resultados reunidos na Tabela 2

mostram que as características *Skewness* e curtoses degradaram muito as taxas de segmentação sucedida e aumentaram as taxas de erro. No caso de cor + média + variância, as diferenças são mínimas em relação a cor + média. A Figura 3 ilustra claramente as semelhanças e diferenças.

Tabela 2: Avaliação numérica da abordagem com o Canal B (em %).

	ME	RAE	TP	TN	FP	FN	P	A	R	ER	FM	NRM
C+M	14.63	13.93	83.40	89.00	11.00	16.60	88.20	86.20	83.40	13.80	84.83	13.80
C+M+V	14.63	13.93	83.40	89.00	12.33	16.60	88.20	86.20	83.40	13.80	84.83	13.80
C+M+S	34.84	35.33	99.06	18.88	90.21	0.94	52.64	54.42	99.06	45.58	68.64	45.58
C+M+C	34.84	35.33	99.06	18.88	90.21	0.94	52.64	54.42	99.06	45.58	68.64	45.58

Figura 4: Visualização gráfica dos resultados da Tabela 2

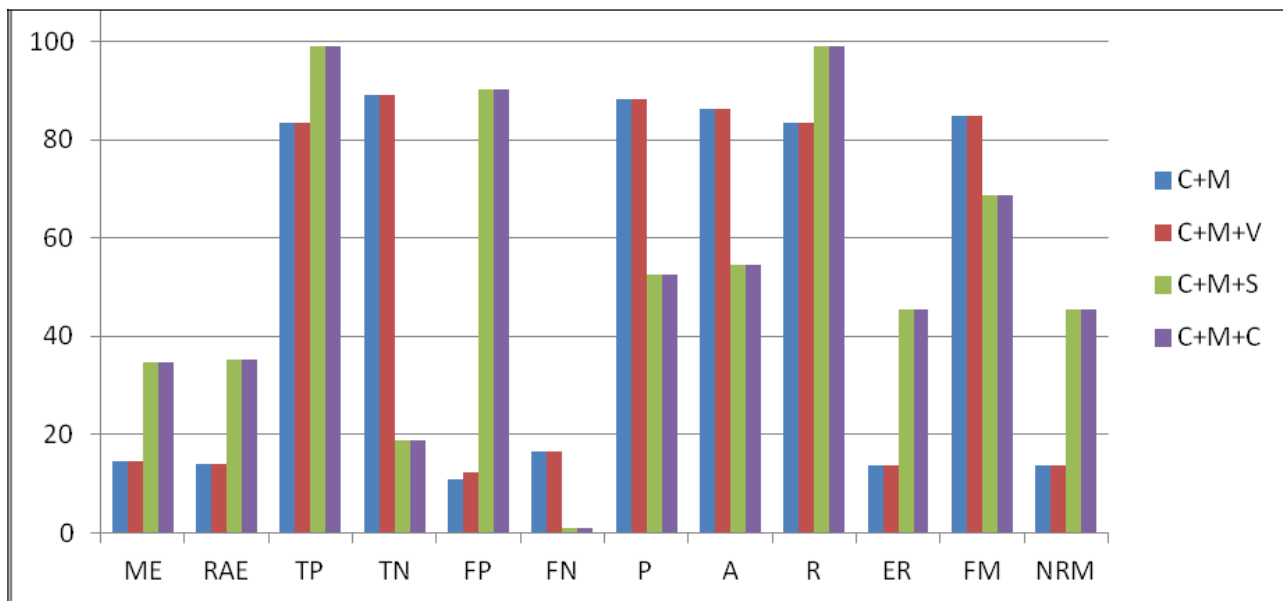
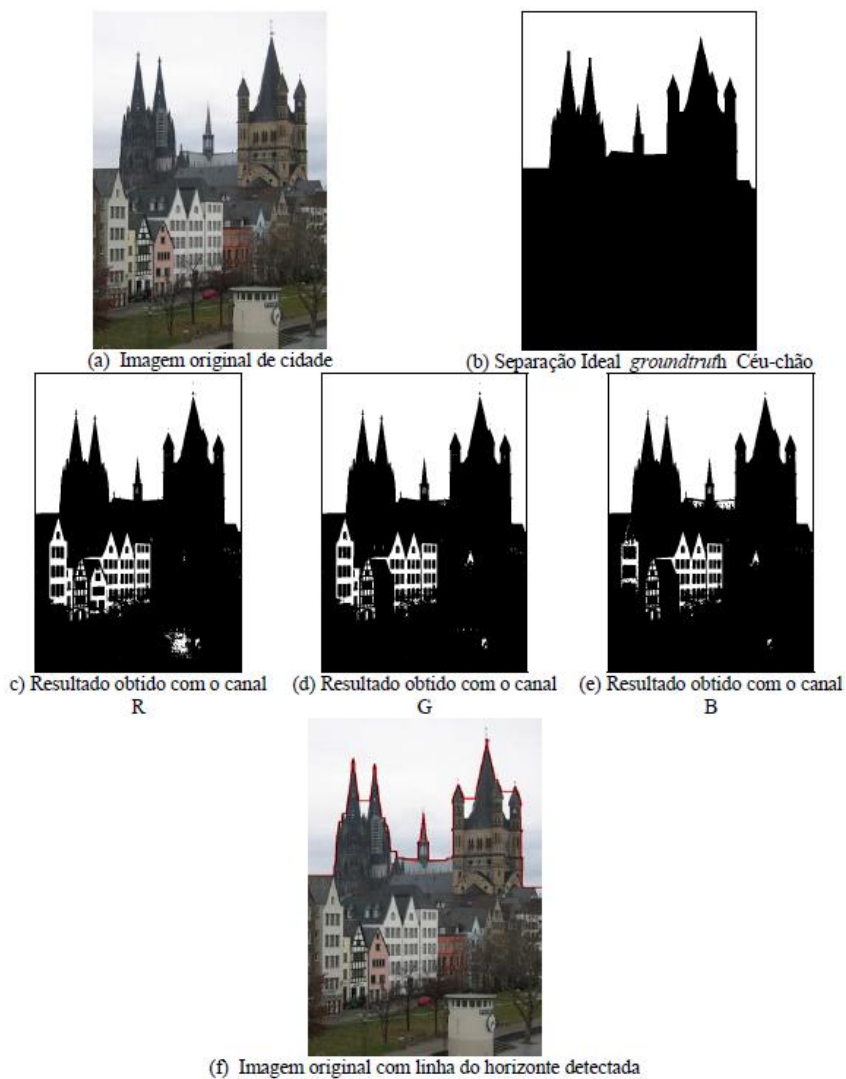


Figura 4: Exemplo de processamento de uma imagem de cidade



As Figuras 4, 5 e 6 ilustram alguns resultados obtidos para três imagens de conteúdos diferentes, uma imagem urbana, outra de paisagem e uma terceira de rodovia. Percebe-se, nas Figuras 4 e 5, que a combinação classificador *KMédio* e algoritmo de binarização de Otsu resultou em segmentações céu-chão semelhantes para os 3 canais R, G e B. No caso da Figura 6 de imagem de rodovia,

percebe-se que o resultado para o canal B é muito mais próxima da separação ideal *groundtruth* céu-chão que para os demais canais R e G.

Nas Figuras 4(f), 5(f) e 6(f) é possível averiguar a precisão dos resultados finais da detecção da linha do horizonte sobreposta em vermelho às imagens originais.

Figura 5: Exemplo de processamento de uma imagem de paisagem

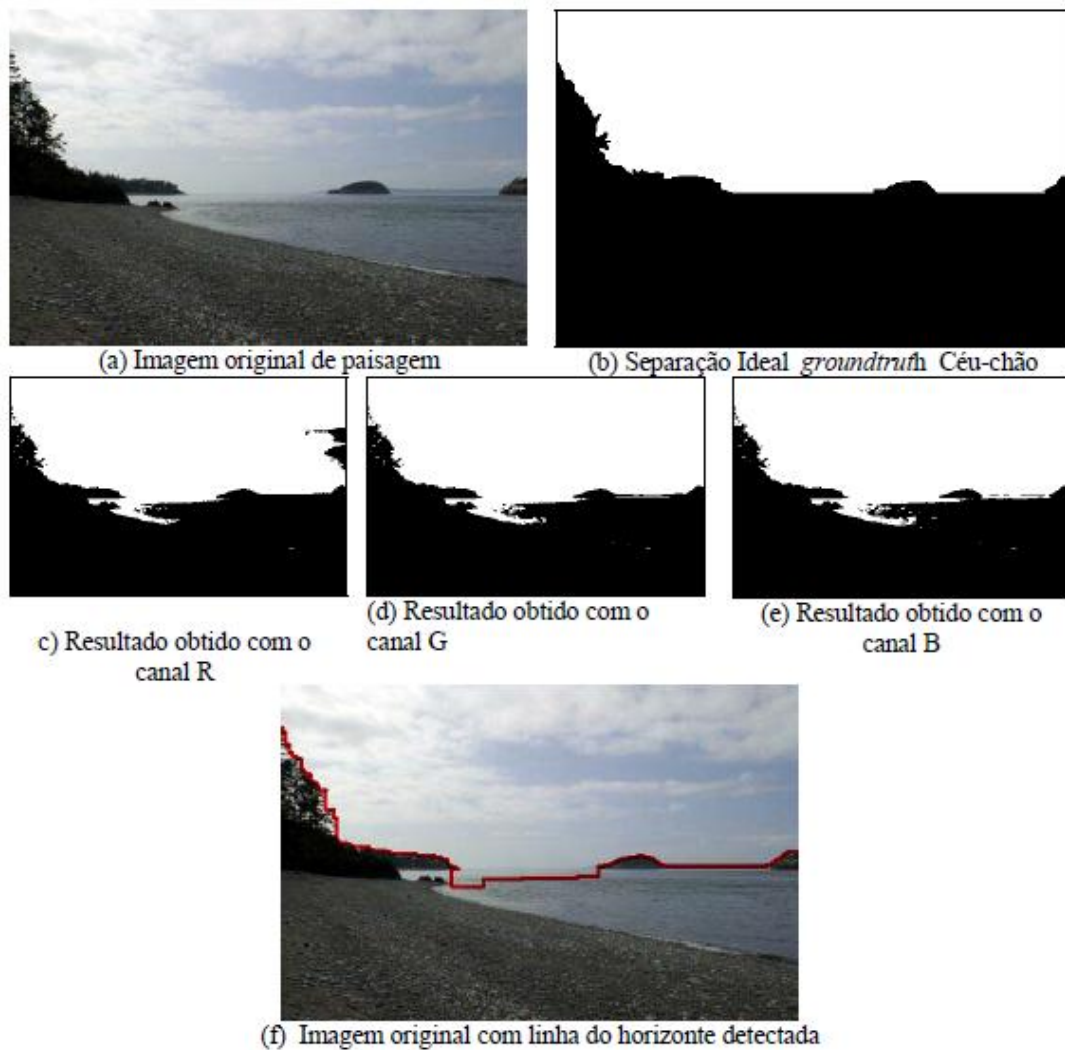
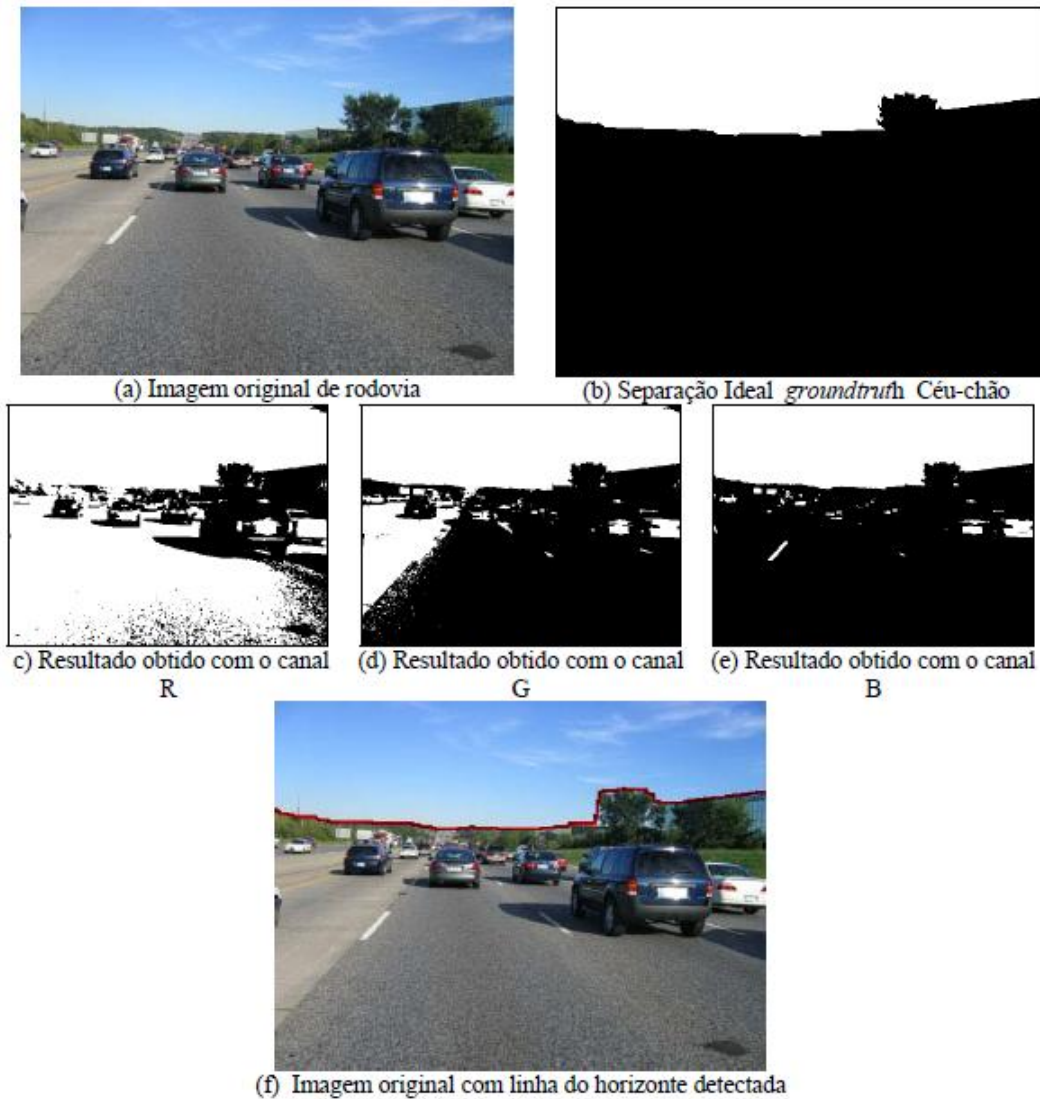


Figura 6: Exemplo de processamento de uma imagem de rodovia.



4. Conclusões

O artigo apresentou uma abordagem de detecção automática de linha do horizonte em imagens coloridas com grandes variabilidades de conteúdos. Vale frisar que essas imagens não tiveram o contraste e brilho previamente corrigidos. O classificador *KMédio* no espaço de cor RGB foi aplicado para segmentar as imagens em duas classes Céu e Chão. A avaliação numérica da eficiência sobre os experimentos realizados a partir de uma base de 44 imagens com a solução ideal *groundtruth* disponível mostraram que a abordagem mostrou-se promissora. E que levando em

consideração todas as métricas e principalmente as métricas *FM* e *NRM*, de 84.83% e 13.8% respectivamente, a combinação canal B + cor + média apareceu como a mais interessante para tratar imagens com diversas variabilidades de conteúdos.

Futuramente pretende-se usar outros espaços de cor considerados mais completos como *HSV* (Matiz, Saturação, Valor) e *YCrCb* (Brilho, Cromo Vermelha, Cromo Azul) assim como o classificador nebuloso Fuzzy *KMédio* acrescentando mais características.

Referências

BIMBO, Del, A., “**Visual Information Retrieval**”, Morgan Kaufmann (1999).

BRAND, J.; MASON, J. “**A comparative assessment of three approaches to pixel-level human skin detection**”, International Conference on Pattern Recognition, Volume 1: pp 1056-1059, 2000.

GATOS, B., Ntirogiannis, K.,Pratikakis, I. , “**ICDAR 2009 Document Image Binarization Contest (DIBCO 2009)**”, International Conference on Document Analysis and Recognition, pp 1375–1382, 2009.

LIAO T. W. “**Clustering of time series data—a survey**”, Pattern Recognition Vol. 38, No 11, pp 1857-1874, 2005.

OTSU N., “**A Threshold Selection Method from Gray-level Histograms**”, IEEE Transactions on Systems, Man and Cybernetics, v. SMC 9, no 1, pp.62-66, 1979.

SANKUR B.; SEZGIN, M., “**Selection of Thresholding methods for non-destructive testing applications**”, International Conference On Image Processing, ICIP2001, pp 764-767, 2001.

Artigo submetido em: 06.08.2015

Artigo aprovado para publicação em: 14.03.2016