

Campos Aleatórios de Markov uma Abordagem para Caracterização Extração de Contornos de Telhados

E.A.S. GALVANIN¹, Departamento de Matemática, Universidade do Estado de Mato Grosso, UNEMAT, 78390-000 Barra do Bugres, MT, Brasil.

A.P. DAL POZ², Departamento de Cartografia, UNESP - Univ Estadual Paulista, 19060-900 Presidente Prudente, SP, Brasil.

Resumo. Este artigo propõe uma metodologia para extração de contornos de telhados utilizando um modelo de Campo Aleatório de Markov (Markov Random Field - MRF). Levando em conta algumas propriedades de telhados e as medidas de alguns atributos (por exemplo, área, retangularidade, ângulos entre eixos principais de objetos) é construída uma função de energia a partir do modelo MRF. O problema de extração de contornos de telhados é formulado a partir de uma estimativa de Maximum a posteriori (MAP), via algoritmo Simulated Annealing (SA). A metodologia proposta foi testada em uma área teste com diferentes complexidades de configurações de objetos presentes na cena.

Palavras-chave. Campo Aleatório de Markov, Extração de Contornos de telhados, Simulated Annealing.

1. Introdução

O MRF, ou Campo Aleatório de Markov, é um modelo que tem atraído muita atenção nos últimos anos. Os modelos MRF têm sido utilizados em aplicações de processamento de imagem de baixo nível, tais como segmentação e restauração de imagem ([4]). No entanto, recentemente vem sendo utilizado em tarefas de análise de imagem de alto nível ([6], [7], [8], [1], [9]). A análise de imagem usando MRF é formulada como um problema de estimação do maximum a posteriori (MAP). Este processo corresponde à resolução de um problema de minimização de energia. Em geral, a função de energia associada com problemas de visão é não-convexo, podendo então ter vários mínimos locais. Assim, a solução pode não corresponder a um mínimo global. Segundo [7], o MRF pode, também, ser definido, sobre grafos e aplicado para o problema de análise de imagem.

¹galvanin@gmail.com

²aluir@fct.unesp.br

2. MRF para Análise de Imagens por Regiões

A formulação de um MRF para problemas de análise de imagem pode ser realizada segundo alguns preceitos, ou seja, parte-se de uma imagem segmentada e constrói-se um grafo de regiões adjacentes (Region Adjacency Graph (RAG)). Cada nó do RAG corresponde a uma região da imagem e dois nós tem conectividade entre eles se as duas regiões correspondentes compartilham de uma mesma fronteira. Em seguida, assume-se que a interpretação do nó, dado o conhecimento específico dos objetos de interesse, e os atributos obtidos da imagem observada, dá-se de acordo com um MRF. Assim, o problema de análise de imagem é então resolvido como um problema de estimativa MAP. Uma das grandes vantagens desta abordagem é a possibilidade de modelar o conhecimento contextual, isto é, as relações entre o objeto de interesse e os demais presentes na cena.

2.1. MRF em estrutura de grafo

Segundo [8], a formulação de um MRF em estrutura de grafos inicia com uma imagem segmentada com n regiões $\{R_1, R_2, \dots, R_n\}$ e o RAG correspondente. Seja $G = \{R, E\}$ um RAG, onde $R = \{R_1, R_2, \dots, R_n\}$ representa o conjunto de nós R_i , $i = 1, 2, \dots, n$ e E denota o conjunto de arestas. Existirá uma aresta entre os nós R_i e R_j se as regiões correspondentes compartilharem, pelo menos parcialmente, de uma mesma fronteira. O sistema de vizinhança em G será representado por

$$\eta = \{\eta R_1, \eta R_2, \dots, \eta R_n\}, \quad (2.1)$$

onde $\eta(R_i)$, $i = 1, \dots, n$ é o conjunto de todos os nós em R vizinhos de R_i . Nota-se que $R_i \in \eta(R_j)$ se $R_j \in \eta(R_i)$.

Uma clique c , neste contexto, é um subconjunto de nós de G (isto é, R) tal que cada par de diferentes nós em c são vizinhos. A coleção de todas as cliques de G com relação ao sistema de vizinhança η é representado como $C(G, \eta)$.

Define-se uma função potencial da clique V_{c_i} envolvendo somente nós de c_i . Assim, cada função desse tipo expressa a forma e o grau de interação (primeira ordem, segunda ordem etc.) que cada nó R_i tem com seus vizinhos. Nesse contexto definiremos $U(x)$ como a função de energia de Gibbs, dada por,

$$U(x) = \sum_{c \in C(G, \eta)} V_c(x^c), \quad (2.2)$$

na qual, $V_c(x^c)$ é a função potencial da clique, sendo x^c o valor das variáveis associadas com os nós pertencentes à clique $c \in C(G, \eta)$. A distribuição de Gibbs, é dada por,

$$P[X = x] = \frac{1}{Z} e^{-U(x)}, \quad (2.3)$$

onde x é uma realização de X e Z é a constante de normalização dada por,

$$Z = \sum \exp^{-U(x)}. \quad (2.4)$$

Seja $X = \{X_1, X_2, \dots, X_n\}$ a família de variáveis aleatórias definida sobre R , onde cada X_i corresponde a R_i . Além disso, assume-se que X_i toma valores em

um espaço amostral finito. Segundo [8], X é denominado um MRF sobre G com relação ao sistema de vizinhança η se e somente se: $P[X = x] > 0$ para todas as realizações de X ; $P[X_i = x_i | X_j = x_j \forall j \neq i] = P[X_i = x_i | X_j = x_j \forall j : R_j \in \eta R_i]$.

Uma das vantagens do modelo de MRF é que geralmente sua função distribuição de probabilidade é dada pela distribuição de Gibbs conforme estabelece o teorema de Hammersley-Clifford [2]. De acordo com [7], assumindo que X tem um número finito de configurações em relação ao espaço amostral S , e que $P[X = x] > 0$, então X é um MRF, com respeito ao sistema de vizinhança η , se e somente se X tem distribuição de Gibbs.

Segundo [10], devido à estrutura na qual as propriedades locais e globais são relacionadas através de cliques, a abordagem baseada no modelo de MRF para análise de imagem fornece vantagens em relação à representação do conhecimento, aprendizado e otimização.

3. Rotulação de Imagem usando MRF

Quando o problema de análise de imagem é restrito à rotulação de regiões a partir de uma imagem segmentada, atribui-se um nó R para cada região. O conjunto de arestas E é tal que o nó R_i está conectado a R_j somente se as regiões correspondentes são espacialmente adjacentes. Cada nó R_i tem um rótulo possível do conjunto $I = \{I_1, I_2, \dots, I_M\}$. Então, o espaço amostral para cada X_i (isto é, a variável aleatória para o nó R_i) será $I = \{I_1, I_2, \dots, I_M\}$. Logo, X_i tomaria um valor do conjunto $\{I_1, I_2, \dots, I_M\}$.

O conjunto de atributos obtidos a partir da imagem de entrada pode ser dado por,

$$F = \{F^c | \forall c \in C(G, \eta)\}, \quad (3.1)$$

onde $F^c = \{F_1^c, \dots, F_q^c\}$, é o conjunto de q atributos medidos sobre a clique c . Esses q atributos seriam os mesmos usados na caracterização do conhecimento a priori κ .

Assume-se agora que a distribuição de probabilidade do vetor aleatório X definido sobre o RAG G , dado o conhecimento a priori κ e o conjunto de atributos F , é um MRF, isto é,

$$P[X = x | F = f, \kappa] = \frac{1}{Z} \exp^{-U(x|f, \kappa)}. \quad (3.2)$$

$$U(x|f, \kappa) = \sum_{c \in C(G, \eta)} V_c(x^c | f^c, \kappa). \quad (3.3)$$

Agora o problema de análise de imagem é resolvido como um problema de estimação do MAP, isto é,

$$x_* = \arg \max P[X = x | F = f, \kappa], \quad (3.4)$$

ou, de forma equivalente,

$$x_* = \arg \min U(x|f, \kappa). \quad (3.5)$$

Para uma determinada imagem, uma análise ótima corresponde à minimização de uma função de energia. No entanto, essa análise ótima depende de como a função de energia é definida.

3.1. Solução MAP

Como foi discutido anteriormente o processo em questão, de análise de imagem, é formulado como um problema de estimativa MAP. Em geral, quando se tem um problema baseado num modelo MRF, a solução procurada geralmente envolve uma estimativa MAP, que é equivalente a um problema de minimização da função de energia dada por,

$$x_* = \arg \min \left[\sum_{\forall c = \text{um nó clique}} V_1^c(x^c | f^c \kappa) + \sum_{\forall c = \text{múltiplos nós cliques}} V_m^c(x^c | f^c \kappa) \right]. \quad (3.6)$$

Em geral, a expressão dentro do colchete na Equação 3.6 possui vários mínimos locais. Assim, é necessária a aplicação de um algoritmo de minimização que possa fornecer um mínimo global. Entretanto, a utilização de um método simples de otimização envolve uma busca combinatorial exaustiva, resultando em uma complexidade exponencial da ordem de L^N , onde L é o número de rótulos e N o número de nós do grafo de adjacência [10].

Todavia, com o intuito de resolver o problema da busca combinatorial, alguns autores têm proposto esquemas de relaxação para encontrar a solução local ótima para o problema da estimativa MAP [11], [5]. Outra solução, descrita na literatura, para reduzir a complexidade exponencial é a utilização do algoritmo Simulated Annealing (SA). Este algoritmo é um procedimento de otimização, que encontra o máximo global da distribuição a posteriori do MRF ou o mínimo da função de energia sem cálculos excessivos [4]. No algoritmo SA, a distribuição de probabilidade condicional é modelada para depender do parâmetro de controle (temperatura),

$$P = [X = x] = \frac{1}{Z(T)} \exp\left(\frac{-U(x)}{T}\right), \quad (3.7)$$

onde

$$Z = \sum_{\text{toda conf. de } x} \exp\left(\frac{-U(x)}{T}\right). \quad (3.8)$$

Um aspecto importante do SA é a escolha da temperatura inicial T_0 , pois a eficácia do algoritmo depende de uma escolha adequada para T_0 . Segundo [10], no algoritmo SA o processo primeiramente "liquefaz" o sistema em uma temperatura alta e, então, gradualmente diminui a temperatura até que o sistema se resfrie e atinja uma configuração ótima. De fato, [12] o apresenta como uma alternativa vantajosa ao princípio de minimização de energia usado para a solução do problema de contorno ativo (snakes), cuja desvantagem está associada à sensibilidade do processo de otimização ao estado inicial. Esses autores ressaltam, ainda, que a literatura não fornece informação sobre a escolha da temperatura inicial, ao contrário, essa escolha envolve tentativa e erro.

4. Caracterização do Conhecimento sobre Contornos de Telhados

Para verificar o desempenho do MRF na análise de imagens por regiões, algumas regiões altas obtidas pela metodologia apresentada por [3] são agora estruturadas segundo um RAG, onde cada nó do RAG corresponde a uma região alta. Essa modelagem deve propiciar a obtenção apenas dos contornos correspondentes aos telhados. Vale ressaltar que, como é necessário o cálculo de vários atributos, deve-se considerar que em uma quadra urbana existe um conjunto de objetos variados, onde a heterogeneidade geralmente é elevada. Isto implica que mais detalhes sobre o objeto de interesse devem ser explorados para auxiliar no seu reconhecimento. Os telhados possuem algumas propriedades de interesse, sendo elas, geométricas, ou seja, as que estão relacionadas com a área, o perímetro, a retangularidade, direção angular etc., e as relações contextuais nas quais o objeto telhado se relaciona com outros objetos presentes na cena (por exemplo, ângulos entre eixos de objetos).

Para definir a clique, inicialmente assumiu-se que os objetos altos (R_i , $i = 1, \dots, n$), imersos num fundo F , são modelados como MRF. A vizinhança η_{R_i} , isto é, das regiões R_j vizinhas de R_i ($i \neq j$), é definida na forma,

$$\eta_{R_i,r} = \{R_j | \text{dist}(R_j, R_i) \leq r\}, \quad (4.1)$$

onde a função dist é dada pela distância euclidiana entre os centros de massa de dois objetos analisados (R_i, R_j) e r é a distância máxima permitida entre R_i e R_j . A Figura 1 mostra um exemplo de vizinhança ($\eta_{R_i,r}$), onde todas as regiões dentro do raio máximo são consideradas vizinhas da região i . Na Figura 1, os pontos representam o centro de massa dos objetos altos.

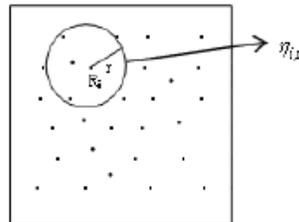


Figura 1: Exemplo de vizinhança.

Uma vez estabelecido o critério de vizinhança $\text{dist}((R_j, R_i) \leq r)$, define-se uma vizinhança individual e, a partir desta definição, tem-se o sistema de vizinhança $\eta = [\eta(R_1), \dots, \eta(R_n)]$. Os conceitos de clique individual C e de coleção de cliques ($C(G, \eta)$) seguem o formalismo apresentado na seção de MRF em estrutura de grafos.

A construção da função de energia $U(I|F, \kappa)$ depende substancialmente do conhecimento a priori sobre as propriedades do objeto telhado. O conhecimento a priori a respeito do objeto de interesse denotado por κ é muito importante na análise de imagem, pois impõe uma forte suposição sobre o que se espera da cena

antes de aplicar o algoritmo para realizar a análise. A caracterização de κ implica em estabelecer valores nominais para os atributos que são considerados importantes para decisão em uma análise. Por exemplo, a área mínima de um edifício pode ser estabelecida como sendo $30m^2$. Já os valores nominais para o conjunto (F) de atributos podem ser medidos no polígono envolvente que contém cada objeto individual e entre polígonos caracterizando a relação contextual entre os objetos. Para tanto é necessário estabelecer um conjunto (F^c) de atributos sobre a clique. O conjunto (F) de atributos é expresso por, $F = \{F^c | \forall c \in C(G, \eta)\}$, onde $F^c = \{F_1^c, \dots, F_q^c\}$, é o conjunto de q atributos medidos sobre a clique individual c . Cabe ressaltar que esses q atributos seriam os mesmos usados na caracterização do conhecimento a priori κ , só que nesse caso eles assumem valores que caracterizam os objetos de interesse.

Os atributos para a clique de primeira ordem utilizados neste trabalho foram a área e a retangularidade. Esses atributos podem ser expressos matematicamente segundo as propriedades geométricas do objeto. Segundo [3] o primeiro atributo baseia-se na área (A) do polígono envolvendo o objeto individual e é definido matematicamente pela fórmula de Gauss dada por:

$$A = \frac{\left| \sum_{i=0}^{n-1} E_{i+1}N_i - \sum_{i=0}^{n-1} E_iN_{i+1} \right|}{2}, \quad (4.2)$$

onde (E, N) correspondem às coordenadas planimétricas de um ponto no referencial do Modelo Digital de Elevação (MDE).

Este atributo permite, por exemplo, que objetos pequenos como caixas d'água, cuja área é relativamente menor em relação aos telhados, possam ser eliminados da análise. Para que isso seja possível, a equação de energia deve penalizar pequenas áreas.

O segundo atributo, retangularidade (R), é obtido através do ângulo formado pela direção principal e secundária do objeto. Para obter o eixo principal de um objeto são calculadas todas as direções dos segmentos de reta de um polígono representando o objeto. Para obter a direção mais freqüente que é correspondente à direção do eixo principal, faz-se uma divisão setorial, particionando o círculo trigonométrico em 24 setores de 15 graus (Figura 2a). Cada setor é representado pelo seu valor angular central (por exemplo, o setor de amplitude $[345^\circ; 360^\circ]$ é representado pelo ângulo $(352,5^\circ)$).

A direção do eixo principal e secundário de um objeto poderia ser representado na forma do histograma da Figura 2b. Esse histograma é obtido após o cálculo da direção e comprimento de cada segmento de reta do polígono que representa o objeto, ou seja para uma dada direção de um dado segmento de reta, verifica-se qual setor que a contém, adicionando a este setor o valor inteiro do comprimento do respectivo segmento de reta, repete-se esse cálculo para todos os segmentos de reta do polígono e assim, identifica-se as direções dos eixos principal e secundário como sendo a primeira e a segunda direção mais freqüente, respectivamente.

Nesse caso, o resultado obtido para a direção do eixo principal e secundário pode ser representado no histograma a seguir onde $7,5^\circ$ é a direção do eixo principal (primeira mais freqüente) e $247,5^\circ$ é a direção do eixo secundário (segunda mais

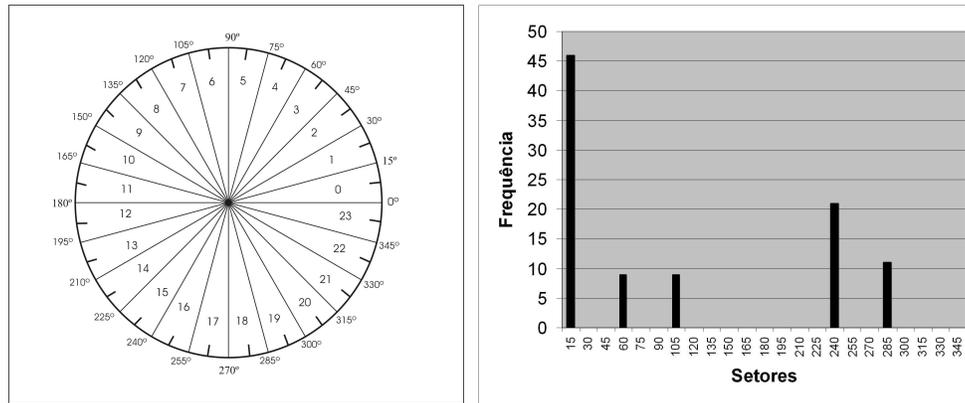


Figura 2: a) Divisão setorial do círculo trigonométrico; b) Histograma de frequência.

frequente). Este resultado foi obtido a partir do contorno de um objeto real, que no caso é um telhado. A retangularidade é expressa matematicamente por,

$$R = |\text{sen } \theta|, \quad (4.3)$$

onde θ é o ângulo entre os eixos principal e secundário.

O atributo R beneficia os objetos com formas geométricas regulares, onde prevalecem os ângulos retos nos vértices do contorno. O valor ótimo para R é 1 (um), sendo este um dos valores a ser incluído no conhecimento κ . O valor de R é 1 (um) em situações ideais, quando $\theta = 90^\circ$ ou $\theta = 270^\circ$.

O terceiro atributo baseia-se em cliques de segunda ordem. Sendo θ_{ij} o ângulo entre as direções principais de dois objetos (R_i, R_j), define-se o seguinte atributo de relacionamento espacial,

$$\Phi(R_i, R_j) = \text{sen}(2 \theta_{ij}). \quad (4.4)$$

Esse atributo possibilita a verificação do paralelismo ou perpendicularismo entre objetos, pois se $\theta_{ij} = 0^\circ$ (objetos com eixos principais paralelos) ou se $\theta_{ij} = 90^\circ$ (objetos com eixos principais perpendiculares), $\Phi(R_i, R_j) = 0$. Portanto, no conhecimento κ deve ser assumido que o valor ótimo para este parâmetro é 0 (zero).

Em relação às cliques, foram utilizadas neste trabalho as de primeira e segunda ordem. A clique de primeira ordem consiste de uma lista de objetos altos com atributos (área e retangularidade) caracterizando os objetos individualmente. Já a clique de segunda ordem é composta de uma lista de objetos, relacionando-se aos pares, caracterizadas por atributos relacionais (no caso o atributo dado pela Equação 4.4).

4.1. Definição da função de energia

A análise de objetos usando a abordagem MRF tem como princípio a minimização da função de energia. Para o problema em questão, espera-se que para um de-

terminado MDE a solução seja ótima, isto é, que seja obtida uma configuração de contornos de telhados, correspondente ao valor mínimo da função de energia. Entretanto, essa análise ótima depende de como a função de energia é definida.

A seguinte Equação de energia foi desenvolvida para a extração de contornos de telhados, a partir de contornos de objetos altos previamente extraídos [3], ficando,

$$\begin{aligned}
 U = & \alpha \sum_{i=1}^n (1 - R_i) + \beta \sum_{i=1}^n \frac{(1 - p_i)}{A_i} + \omega \sum_{i=1}^n \sum_{j \in (G, \eta)} p_i p_j |\text{sen}(2 \theta_{ij})| \\
 & + \gamma \sum_{i=1}^n [p_i \ln p_i + (1 - p_i) \ln(1 - p_i)].
 \end{aligned}
 \tag{4.5}$$

onde α , β , ω e γ são pesos que dão a importância relativa para cada termo das funções de energia; R_i é a medida de retangularidade do objeto R ; A_i é a área do objeto R_i ; p_i (ou p_j) é uma medida individual de compatibilidade de R_i (ou R_j) com um contorno de telhado; θ_{ij} é o ângulo entre as direções dominantes dos objetos R_i e R_j .

Minimizar a função de energia U (equação 4.5) implica minimizar simultaneamente os quatro termos de energia de U . No final do processo de minimização, isto é, quando U for mínimo, obtém-se uma configuração ótima dos contornos que são telhados de edifícios. O valor final de p_i para contornos de telhados é um, enquanto que para os outros objetos é zero. O algoritmo de otimização utilizado foi o SA, que é eficiente na obtenção do mínimo global, mesmo quando a função de energia possui mínimos locais.

5. Resultados

Nesta seção são apresentados e analisados os resultados da metodologia proposta para a extração de contornos de telhados a partir da caracterização do conhecimento de telhados. Para a comparação visual dos resultados obtidos na Figura 3b foi utilizada uma imagem de intensidade obtida por perfilamento a laser, Figura 3a, fornecida pelo Instituto de Tecnologia para o Desenvolvimento (LACTEC).

A Figura 3(b) mostra que um dos telhados foi extraído mesmo possuindo um porte bem menor que o outro telhado, dada a forma que a Equação de energia foi definida (Equação 4.5) objetos de área relativamente menor podem ser descartados. Entretanto, fica evidente que a situação é bem favorável para os atributos de relacionamento espacial e de retangularidade, beneficiando a extração dos dois telhados existentes na cena. O modelo MRF possibilitou a utilização de injunções espaciais, as quais permitiram modelar melhor a cena. As injunções espaciais podem auxiliar na discriminação dos objetos de interesse em várias situações, mesmo que o atributo de retangularidade não seja bom.

6. Conclusões

Neste trabalho foi apresentada uma metodologia de caracterização do conhecimento de contornos de telhados de edifícios em um MDE obtido a partir de dados de

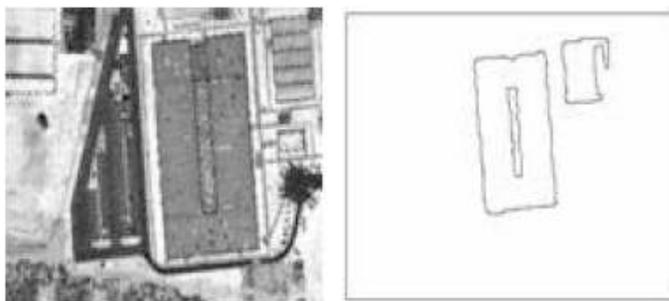


Figura 3: (a) Imagem intensidade referente à área teste; (b) resultado da extração usando a equação de energia via algoritmo SA.

varredura a laser, bem como o resultado obtido com a metodologia proposta. A contribuição principal deste trabalho está relacionada com a exploração de relações espaciais entre os edifícios na cena, o que foi viabilizado pelo modelo MRF.

Agradecimentos Os autores são gratos ao Instituto LACTEC (pela concessão de dados laser).

Abstract. This paper proposes a methodology for automatic extraction of building roof contours by Markov Random Field Model (MRF). Taking into account some roof properties and some feature measurements (e. g., area, rectangularity, and angles between principal axes of the roofs), an energy function was developed based on the Markov Random Field (MRF) model. The problem of building roof contour extraction is formulated as the Maximum a Posteriori (MAP) estimation by Simulated Annealing (SA) algorithm. The proposed methodology was tested in one test areas with different object configuration complexities.

Referências

- [1] H. Andersen, S. Reutebuch, G. Schreuder, Bayesian object recognition for the analysis of complex forest scenes, In: “Airbone Laser Scanner Data. ISPRS, Symposium Photogrammetric Computer Vision”, Áustria, 2002.
- [2] J. Besag. Spatial interaction and the analysis of lattice systems. *J. Royal Statistical Association, Series B* **36** (1974), 192-236.
- [3] E.A.S. Galvanin, A.P. Dal Poz, A.D.P. Souza, Segmentação de dados de perfilamento a laser em áreas urbanas utilizando uma abordagem bayesiana, *Boletim de Ciências Geodésicas*, **13** (2007), 76-90.
- [4] S. Geman, D. Geman, Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1984), 721-741.

- [5] R.A. Hummel, S.W. Zucker, On the foundation of relaxation labeling process. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1983) 267-287.
- [6] I.Y. Kim, H.S. Yang, An integrated approach for scene understanding based on Markov Random Field, *Pattern Recognition*, (1995), 1887-1897.
- [7] R. Kinderman, J.L. Snell, "Markov Random Fields and their Applications", Providence, R.I., American Mathematical Society, 1980.
- [8] S.K. Kopparapu, U.B. Desai, "Bayesian Approach to Image Interpretation", 2001.
- [9] S.Z. Li, "Markov Random Field Modeling in Image Analysis", Computer Science Workbench, 3rd ed., 2009, XXII, 362p.
- [10] J.A. Modestino, J. Zhang, A Markov Random Field model based approach to image interpretation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6** (1992), 606-615.
- [11] A. Rosenfeld, R.A. Hummel, S.W. Zucker, Scene labeling by relaxation operations, *IEEE Trans. Syst. Man Cybern.* **SMC-6** (1976), 420-433.
- [12] J.C. Trinder, U. Maulik, S. Bandyopadhyay, Semi-automated feature extraction using simulated annealing, *International Archives for Photogrammetry and Remote Sensing*, (2000), 905-911.
- [13] G. Winkler, "Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction", Series: Stochastic Modelling and Applied Probability , vol. 27, 2nd ed., 2006, XVI, 389p.