Anna Fabijańska*

Graph Based Image Segmentation

1. Introduction

Image segmentation is one of the fundamental problems in machine vision. In general it aims at extracting objects of interest from background. Because this step is crucial for almost every application of vision systems, problem of image segmentation has been widely studied over the years and a very large number of segmentation techniques have already been proposed.

Recent research on image segmentation has seen an increasing interest in graph based techniques which extract objects from background by partitioning graphs into subgraphs. The most representative methods for graph based segmentation are [1]: spectral graph partitioning and combinatorial graph cuts. The first group of methods use the eigenvectors of the graph Laplacian (or its variants) to partition the graph [2]. Whereas combinatorial graph cuts try to define subgraphs by solving min-cut/max-flow problem [3, 4].

Spectral graph partitioning methods are of NP-complexity therefore they often apply some relaxing techniques to find an approximate cut of a graph. These methods however will not be regarded in this paper. For further details the reader is referred to [1, 2].

The main attention of this paper is focused on these image segmentation methods which use combinatorial graph cuts. Specifically, the efficient method of min-cut/max-flow segmentation proposed by Boykov and Jolly in [3] is investigated. The method divides image into subregions by computing a global optimum among all segmentations satisfying some hard constrains imposed for object and background. The optimum is computed in term of minimization of energy which incorporates some regional and boundary conditions.

Although the considered method raised an interest and popularity it has not been fully investigated yet. Incorporating different conditions into energy being minimized leads to different results. Therefore this paper considers what energy kind of can be minimized by graph cuts and how incorporated regional and boundary terms influence properties of segmented image.

---

* Technical University of Lodz, Lodz, Poland

This following part of this paper is organized as follows. Firstly in Section 2 some basic background on graph theory is given. Next, in Section 3 graph representation of an image is explained. In Section 4 the idea of graph based segmentation is explained. The glance at min-cut/max-flow image segmentation is given in Section 5. The results of image segmentation obtained for various images using different energy functions are presented and discussed in Section 6. Finally, Section 7 concludes the paper.

2. Background on graph theory

Graph is a data structure consisting of set of vertices (or nodes) and set of edges connecting the vertices. In more formal sense graph G is often defined as an ordered pair $G = (V, E)$ comprising set V of vertices together with set E of edges. Each edge e_{ij} relates two vertices and is described by a set $\{v_i, v_j\}$ of two corresponding elements from set V. If pairs of vertices $\{v_i, v_j\}$ are unordered edges have no orientation the graph is called undirected or simple. Additionally, when every node is connected with all other nodes the graph is called fully connected [5].

Exemplary undirected graph is presented in Figure 1a. Set of vertices and set of edges corresponding with the graph equal respectively V = {A, B, C, D, E, F} and E = {{A,B}, {A,C}, {A,D}, {B,C}, {B,E}, {D,E}, {D,F}}.
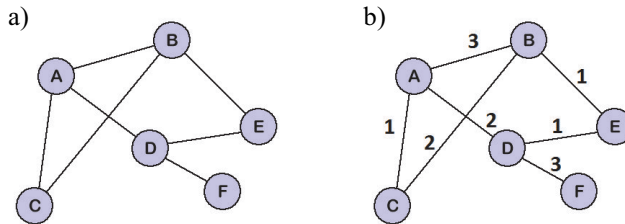

Fig. 1. Exemplary graphs; (a) simple, undirected graph; (b) weighted undirected graph

When weight w_{ij} is associated with every edge e_{ij}, the graph is called a weighted graph. Weights are usually real numbers which express some property of connection between the nodes. Exemplary weighted graph is presented in Figure 1b.

A weighted directed graph with two terminal nodes S (source) and T (sink) can be considered as a flow network with edge capacities determined by its weights. An edge can receive flow not exceeding its capacity. A cut (S, T) of a directed graph is a set of edges $C@E$ such that the two terminals become separated on the induced graph $G' = (V, E/C)$. Minimum cut (min-cut) is a cut of minimum total capacity. According to max-flow min-cut theorem the minimal cut is equal to maximum flow that can be passed from the source S to sink T. This concept is presented in Figure 2. The maximum flow which can be send from S to T range 7 and is equal to total capacity of minimal cut C = {{S,A},{S,B}}.
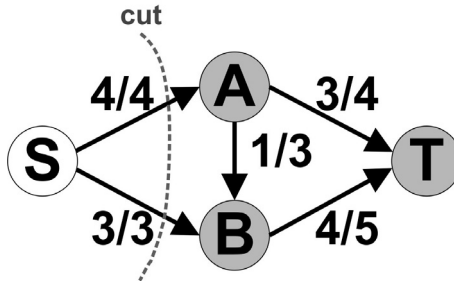
Fig. 2. An idea of min-cut/max-flow theorem

3. Graph based representation of an image

Graph based approaches to image segmentation consider image as a weighted graph where pixels are represented by nodes $v_i \in V$ and each edge $e_{ij} = \{v_i, v_j\} \in E$ has the corresponding nonnegative weight w_{ij}. Weights describe similarity (or dissimilarity) between the neighboring pixels.

Regarding graph based image representation explained above image segmentation is then simply partitioning V into two disjoint sets A and B where $A \cup B = V$ and $A \cap B = \emptyset$. The partitioning is performed according to some criterion and aims at removing edges that connect parts A and B (see Fig. 3).
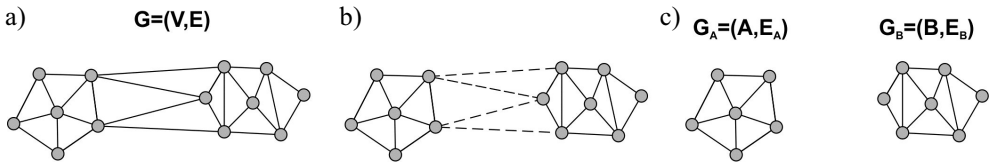
Fig. 3. Idea of graph partitioning; (a) input graph $G = (V, E)$; (b) edges to be removed are denoted by dashed line; (c) disjoint subgraphs $G_A = (A, E_A)$ and $G_B = (B, E_B)$ of graph G

The main problems to be solved during graph-based image segmentation is how to choose edges to be removed in order to divide graph into pieces. Boykov and Jolly proposed method which determines these edges as minimum cut of a planar graph composed from consecutive pixels [3]. The method is briefly characterized in the next section.

4. Glance at min-cut/max-flow image segmentation

4.1. Graph construction

In the regarded method an image is represented by a weighted undirected graph $G = (V, E)$ where nodes correspond with pixels p of an image P. Additionally, in the graph there are

also two special nodes: an object terminal S (called source) and a background terminal T (called sink). Therefore:

$$V = P \cup \{S, T\} \tag{1}$$

The set E consists of two types of undirected edges: *n-links* which connect neighboring pixels and *t-links* which connect pixels with terminals. Every pixel has up to four *n-links* to its (spatially) nearest neighbors and two *t-links*: $\{p, S\}$ and $\{p,T\}$ connecting it to source and sink respectively. Weights $B_{\{pq\}}$ assigned to *n-links* represent boundary term and describe similarity between the neighboring nodes p and q. Specifically, the higher the weight – the higher similarity between pixels; values close to 0 correspond with very low similarities. Weights $R_p(\cdot)$ assigned to *t-links* represent regional term and define the individual penalties $R_p(\text{obj})$ and $R_p(\text{bkg})$ for assigning pixel p to object and background respectively. Exemplary graph obtained for a 3×3 image is presented in Figure 4. The weights of edges are given in Table 1. Symbols O and B used in Table 1 denote object and background respectively.
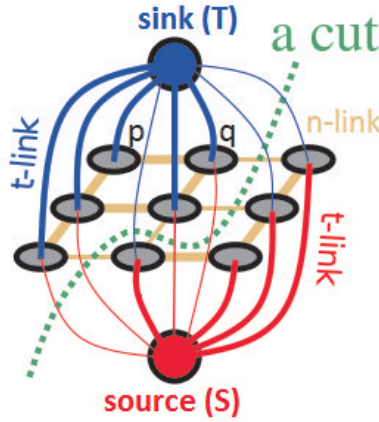
Fig. 4. Exemplary graph obtained for a 3×3 image [3]

Table 1
Weights for *n-links* and *t-links*

Edge	Weight	For
$\{p, q\}$	$B_{\{pq\}}$	$\{p, q\} \in N$
$\{p, S\}$	$\lambda \cdot R_p\,(\text{bkg})$	$p \in P, p \notin O \cup B$
	K	$p \in O$
	0	$p \in B$
$\{p, T\}$	$\lambda \cdot R_p\,(\text{obj})$	$p \in P, p \notin O \cup B$
	0	$p \in O$
	K	$p \in B$

Value of K is defined by following equation:

$$K = 1 + \max_{p \in P} \sum_{q:\{p,q\}} B_{\{p,q\}} \qquad (2)$$

and λ is a scaling factor indicating the relative importance of the regional term versus the boundary term.

4.2. Image segmentation

Having the graph image representation as defined in the previous subsection the border between the object and the background (optimal in terms of properties build into the edge weights) is a minimum cost cut on the graph. The minimum cut is determined via minimization of energy function E given by Equation (3) .

$$E(A) = \lambda \cdot R(A) + B(A) \qquad (3)$$

where $A = (A_1, A_2, ..., A_{|P|})$ is segmentation vector specifying the assignment of pixel p in P either to the background or to the object. Specifically, $A_i = 1$ if p belongs to object; otherwise $A_i = -1$. Additionally:

$$R(A) = \sum_{p \in P} R_p(A_p) \qquad (4)$$

$$B(A) = \sum_{\{p,q\}} B_{\{p,q\}} \cdot \delta(A_p, A_q) \qquad (5)$$

where:

$$\delta(A_p, A_q) = \begin{cases} 1 & A_p \neq A_q \\ 0 & A_p = A_q \end{cases} \qquad (6)$$

According to min-cut/max-flow theorem segmentation vector A that minimizes energy function given by Equation (3) is equivalent to maximum flow that can be send via graph defined in Subection 4.1 from terminal S to terminal T. Border between the object and the background is indicated by edges that gets saturated when the maximal flow is send between S and T.

To find maximal flow Boykov and Kolmogorov proposed new end efficient algorithm based on augmenting paths [4]. However, they didn't define clear and unequivocal criterions for assigning weights to *t-links* and *n-links*. In the next sections results of using different boundary and regional conditions are analyzed.

5. Input data

8-bit monochromatic images of different character and levels of details are regarded in this work. The spatial resolution of each image is 256×256 pixels. The images and their names are presented in Figure 5.

(a) starfish   (b) plane   (c) cars

(d) texture   (e) cottage   (f) pentagon
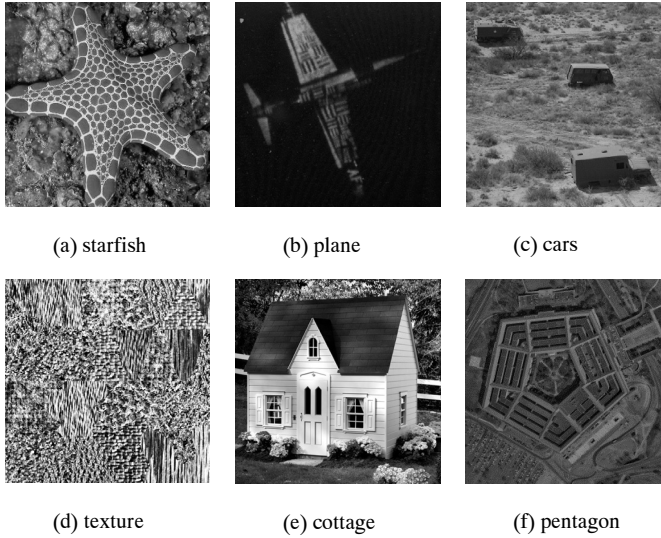
Fig. 5. Test images used to test the performance of graph cut algorithm; (a) starfish; (b) plane; (c) cars; (d) texture; (e) cottage; (f) pentagon

6. Tests and results

In this section the influence of boundary and regional conditions on image segmentation using graph cut approach is presented on exemplary images (see Sec. 5).

Before the processing, regions belonging to the foreground and the background were indicated manually in order to impose some hard constrains on image segmentation. Input images with frames indicating the foreground and the background are presented in Figure 6a and Figure 7a (first columns). Specifically, red color represents regions which must be included into the background while green color indicates images regions contained in the foreground. Properties of these regions were next used to qualify the remaining pixels into these areas.

For describing the boundary conditions two functions given by Equations (7) and (8) were tested. As they are to represent the similarity between the neighboring nodes p and q their values decrease with increasing difference between intensities I_p and I_q of the neighboring nodes.

$$B_{\{p,q\}} = 255 - |I_p - I_q| \tag{7}$$

$$B_{\{p,q\}} = e^{-|I_p - I_q|} \tag{8}$$

While regarding boundary conditions four cases were investigated. Firstly (Case 1), only difference between the intensity I_p of the node p and the average intensities of background $\bar{I}_B$ and foreground $\bar{I}_O$ were considered and weights were determined as described by Equation (9).

$$\begin{cases} R(\text{obj}) = 1 - (I_p - \bar{I}_B)/\bar{I}_B \\ R(\text{bkg}) = 1 - (I_p - \bar{I}_O)/\bar{I}_O \end{cases} \tag{9}$$

Next (Case 2), the probabilities of each intensity to belong to foreground and background were used for describing boundary conditions in accordance with the following equation:

$$\begin{cases} R(\text{obj}) = Pr(I_p \,|\, O) \\ R(\text{bkg}) = Pr(I_p \,|\, B) \end{cases} \tag{10}$$

where: $Pr(I_p|O)$ and $Pr(I_p|B)$ denote probabilities of intensity I_p to belong to the foreground and the background respectively. The probabilities were determined based on intensity distribution in every region.

This was followed by extending the regional terms by gradient properties (Case 3). Specifically, average gradient magnitude and average gradient direction in foreground and background supplemented the average region intensity regarded in the Case 1 (see Eq. (11)).

$$\begin{cases} R(\text{obj}) = 1 - \dfrac{(I_p - \bar{I}_O)(\partial I_p - \bar{\partial} I_O)(\varphi(\partial I_p) - \bar{\varphi}(\partial I_O))}{\bar{I}_O \bar{\partial} I_O \bar{\varphi}(\partial I_O)} \\[4mm] R(\text{bkg}) = 1 - \dfrac{(I_p - \bar{I}_B)(\partial I_p - \bar{\partial} I_B)(\varphi(\partial I_p) - \bar{\varphi}(\partial I_B))}{\bar{I}_B \bar{\partial} I_B \bar{\varphi}(\partial I_B)} \end{cases} \tag{11}$$

where:

$\bar{I}_O, \bar{I}_B$ – average intensity in foreground and background region,

I_p – intensity of node p,

$\bar{\partial} I_O, \bar{\partial} I_B$ – average gradient magnitude in foreground and background region,

∂I_p – gradient magnitude in node p,

$\bar{\varphi}(\partial I_O), \bar{\varphi}(\partial I_B)$ – average gradient direction in foreground and background region,

$\varphi(\partial I_p)$ – gradient direction in node p.

Finally (Case 4), probabilities of each pixel to belong to foreground and background were investigated based on analysis of image intensity distribution, gradient magnitude distribution and gradient direction distribution in accordance with Equation (12).

$$\begin{cases} R(\text{obj}) = Pr(I_p \,|\, O) \cdot Pr(\partial I_p \,|\, O) \cdot Pr(\varphi(\partial I_p) \,|\, O) \\ R(\text{bkg}) = Pr(I_p \,|\, B) \cdot Pr(\partial I_p \,|\, B) \cdot Pr(\varphi(\partial I_p) \,|\, B) \end{cases} \quad (12)$$

where: $Pr(I_p|O)$ and $Pr(I_p|B)$ denote probabilities of intensity I_p to belong to the foreground and the background respectively, $Pr(\partial I_p|O)$ and $Pr(\partial I_p|B)$ denote probabilities of gradient magnitude ∂I_p to belong to the foreground and the background respectively and $Pr(\varphi(\partial I_p)|O)$ and $Pr(\varphi(\partial Ip)|B)$ are probabilities of gradient direction $\varphi(\partial I_p)$ to belong to the foreground and the background respectively. The probabilities were determined based on the corresponding property distribution in every region.

Results of min-cut/max-flow segmentation applied to exemplary images are shown in Figures 6 and 7. Specifically, Figure 6 corresponds to results obtained using Equation (7) for describing boundary conditions and consecutively Equations (9)–(12) (Cases 1–4 respectively) for describing regional conditions. In Figure 7 results obtained using Equation (8) for boundary conditions and Equations (9)–(12) for regional conditions are shown. In both Figures the first column presents original images with regional constrains imposed on the background and the foreground. The following columns present results obtained in Cases 1–4. Case id is indicated in a column caption.

Results presented in Figures 6 and 7 clearly show that selection of functions describing the boundary and regional relations in the image significantly influences the results of image segmentation. Firstly, it should be stated, that using exponential function for describing boundary conditions penalizes significantly discontinuities between pixels of similar intensities. The effect is clearly seen in Figure 7 when visibly similar neighboring pixels are included into the different regions. This make the function insufficient for segmentation of regions covered by highly varied texture (see Fig. 7, images: *starfish*, *texture*). In this case linear function provides significantly better results (see Fig. 6) and successfully segments objects even of very diverse patterned texture.

When regarding regional conditions it should be stated that uniform objects can be successfully extracted from the scene via graph cut, based only on mean intensities of regions indicated as the background and the foreground (see Eq. (9)). This was in case of *plane* image (Fig. 5, Case 1). Investigating probabilities of pixel intensities to belong to the certain regions was satisfactory for segmentation of moderately differentiated textures (see Fig. 6, images: *starfish*, *tanks*, *pentagon*, Case 2). Finally, in order to obtain satisfactory segmentation of objects characterized by extremely diverse texture pattern it was necessary to include gradient distribution information into function describing boundary properties (see Fig. 6, image: *texture*, Case 4).
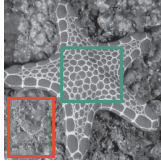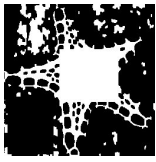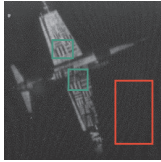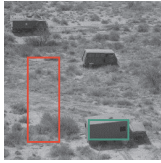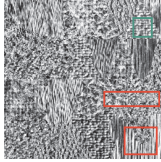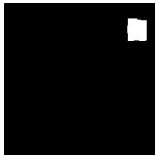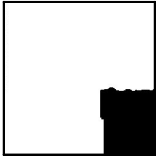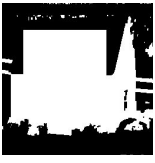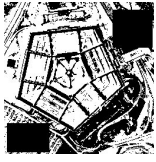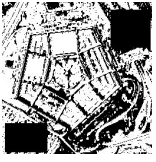
Constrains	Case 1	Case 2	Case 3	Case 4
(a)	**(b)**	**(c)**	**(d)**	**(e)**

Fig. 6. Results of graph min-cut/max-flow segmentation applied to exemplary images. For boundary conditions linear function given by Equation (7) was applied
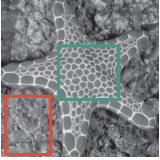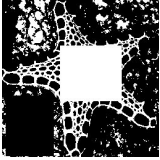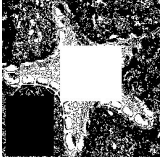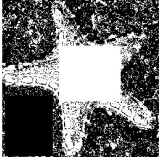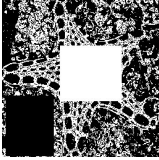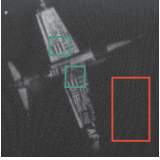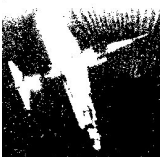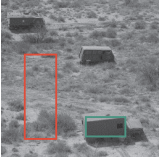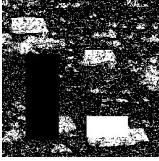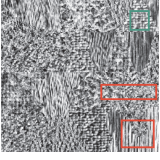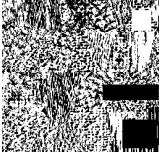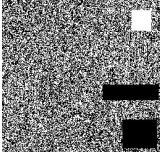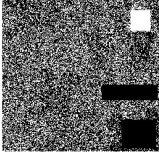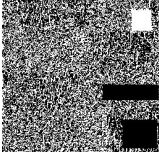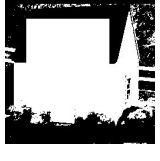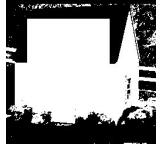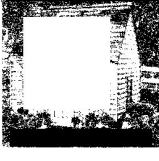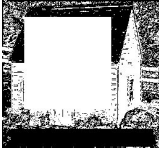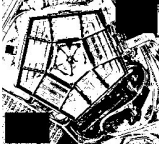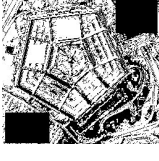
Constrains	Case 1	Case 2	Case 3	Case 4
(a)	(b)	(c)	(d)	(e)

Fig. 7. Results of graph min-cut/max-flow segmentation applied to exemplary images.
For boundary conditions exponential function given by Equation (8) was applied

7. Conclusions

In this paper graph min-cut/max-flow segmentation algorithm was investigated. Although the method has recently gained an increasing interest of the researchers, it lacks clear and unequivocal information how to describe boundary and regional conditions used during image segmentation.

Therefore, problem of boundary and regional penalties selection was investigated in this paper. Functions for describing these both terms were proposed and their influence on image segmentation results were investigated. The obtained results can be then treated as a guidelines for development of new boundary and regional constrains for graph min-cut/max-flow segmentation.

Acknowledgements

References

[1] Wang J., *Graph Based Image Segmentation. A Modern Approach*. VDM Verlag, Germany, 2008.

[2] Shi J., Malik J., *Normalized Cuts and Image Segmentation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, nr 8, t. 22, 2000, 888–905.

[3] Boykov Y., Jolly M.-P., *Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images*. International Conference on Computer Vision, t. 1, 2001, 105–112.

[4] Boykov Y., Kolmogorov V., *An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision*. IEEE Transactions on Pattern Analysis and Machine Intelligence, nr 9, t. 26, 2004, 1124–1137.

[5] Diestel R., *Graph Theory*. Springer, Berlin, Germany, 2010.