Partial optimal labeling search for a NP-hard subclass of (max,+) problems

Ivan Kovtun

International Research and Training Center of Information Technologies and Systems, Kiev, Ukraine, kovtun@image.kiev.ua
Dresden University of Technology, kovtun@ics.inf.tu-dresden.de

Abstract. Optimal labeling problems are NP-hard in many practically important cases. Sufficient conditions for optimal label detection in every pixel are formulated. Knowing the values of the optimal labeling in some pixels, as a result of applying the proposed algorithm, allows to decrease the complexity of the original problem essentially.

1 Introduction

Labeling problems play a significant role in computer vision. In the present paper labeling problems are formulated as (max,+) problems. A subconvex subclass of (max,+) problems is introduced. Many image recognition problems lead to (max,+) problems. For instance, energy minimization, image segmentation by texture features, three-dimensional object reconstruction on the basis of stereoimages, noisy image restoration and other.

For the first time the (max,+) problem was formulated in the paper [8] in 1976. In the papers [8] and [11] an algorithm based on the substitution of the (max,+) problem by an auxiliary linear programming problem is proposed. The algorithm takes the decision itself whether the solution of the auxiliary linear programming problem leads to the solution of the (max,+) problem or not. Thus, the algorithm returns either the solution of (max,+) problem or the answer ”no answer”. In [9] and [10] a subclass of the solvable (max,+) problems is determined and it is shown that for this subclass the algorithm proposed in [8,11] always finds the solution (the answer ”no answer” never appears). These problems belong to the subconvex subclass of (max,+) problems.

A new branch of algorithms for solving (max,+) problems was formed in works [3,4,6]. These algorithms are based on a reduction of the initial problem to a min-cut problem. At that the class of solvable (max,+) problems was not expanded: only a subset of subconvex (max,+) problems is solvable by proposed algorithms.

Thus, a rather well investigated subclass of the (max,+) problems that can be solved in polynomial time arises. Namely, a subset of subconvex (max,+) problems. At the same time there are practically significant subclasses of the (max,+) problems that are known to be NP-hard. Therefore, a number of investigations were devoted to a searching for approximative algorithms [5,2,7,1].

However, the exact solution of the problem in some separately taken pixels is of interest as well.

This paper presents sufficient conditions for making decision about the optimal labeling in each pixel individually. At that the special label "no label" is allowed. The proposed algorithm is applied for a NP-hard subclasses of (max,+) problems namely the Potts model. It can be also generalized to arbitrary (max,+) problems.

2 The basic definitions

Basic definitions are introduced in this section. Later we operate with such notions as ***a vision field, a pixel, a labeling, a label set, a structure of the vision field, an order of the structure and neighboring pixels***. After the definition of ***a labeling quality*** the (max,+) problem is formulated in general form as the problem of searching for the labeling with an optimal quality.

Let ***a vision field*** T be an arbitrary finite set. The elements of the vision field are called ***pixels***. One of the most frequently encountered examples of a vision field is a rectangular area of a two-dimensional integer lattice $\{(i,j)\,|0 \le i < I, 0 \le j < J\}$. Let ***a labeling*** of the vision field T be a function $k_T : T \to \{1, 2, \dots, l\}$. The set $\{1, 2, \dots, l\}$ is called ***a label set*** and denoted by the symbol L. A restriction of this function on a subset $\tau \subseteq T$ of the vision field is denoted by k_τ ($k_\tau : \tau \to L$), and the value of the function k_T in the pixel t is denoted by k_t. Let ***a structure of the vision field*** T be a set $\Im \subseteq 2^T$ of subsets of the vision field T. Note that $\Im$ does not necessarily contain all subsets of the vision field. Let ***the order of the structure*** be a maximum cardinality of the elements of the structure $\Im$, i.e. $\max\limits_{\tau\in\Im} |\tau|$. Usually but not necessarily the order of the structure is two. Pixels t and t' are called ***neighboring*** according to the structure $\Im$ if there exists a subset $\tau \in \Im$ that contains both pixels $\{t, t'\} \subseteq \tau$. Let us denote the set of all labelings of the part τ of the vision field by $L^\tau = \{k_\tau | k_\tau : \tau \to L\}$. A function $g_\tau : L^\tau \to R$ is given for every subset $\tau \in \Im$ of the structure. This function assigns a real number to every labeling $k_\tau : \tau \to L$. Let ***the quality of the labeling*** $k_T : T \to L$ be the number

$$Q(k_T) = \sum_{\tau\in\Im} g_\tau(k_\tau). \tag{1}$$

The (max,+) problem consists in maximization of the quality function $Q()$ and determination of the appropriate labeling $k_T^* = \arg\max\limits_{k_T} Q(k_T)$.

3 A known polynomially solvable subclass of (max,+) problems

Let us suppose that the label set L is a completely ordered set: $1 < 2 < \cdots < l$. A partial ordering can be defined on the set of label pairs $(l, l') \in L \times L$. Namely, for every two label pairs (r, r') and (l, l') their maximum $(r, r') \vee (l, l') = (r \vee l, r' \vee l')$ and minimum $(r, r') \wedge (l, l') = (r \wedge l, r' \wedge l')$ are defined in a natural way. In the same way a partial ordering can be defined on the set of all labelings. For each pair of labelings k_T and k'_T we denote by $k_T \bigvee k'_T$ their maximum and by $k_T \bigwedge k'_T$ their minimum.

A function $Q : L^T \to R$ is called ***subconvex*** if the following condition is fulfilled for arbitrary labelings k_T and k'_T:

$$Q(k_T) + Q(k'_T) \leq Q(k_T \bigvee k'_T) + Q(k_T \bigwedge k'_T). \tag{2}$$

We call a (max,+) problem subconvex if its quality function is subconvex.

An equivalent definition of subconvex function $Q : L^T \to R$ is based on the notion of discrete derivation: $\forall k_T : T \to L, k_t < l$:

$$Q'_t(k_1, \ldots, k_t, \ldots, k_n) = Q(k_1, \ldots, k_t + 1, \ldots, k_n) - Q(k_1, \ldots, k_t, \ldots, k_n).$$

The function $Q : L^T \to R$ is ***subconvex*** if and only if its second derivative $Q''_{tt'}(k_T) \geq 0$ is not negative for every two neighboring pixels t and t' $(t \neq t')$ and an arbitrary labeling $k_T : T \to L$ $(k_t < l, k_{t'} < l)$.

All solvable (max,+) problems known so far are subconvex. It is already proved that a subconvex (max,+) problem is solvable in polynomial time if it fulfills one of the following conditions:

1. The order of the vision field structure is two [8,11,9,10,1,3,4].
   Then the quality function (1) takes the following form
   $$Q(k_T) = \sum_{\{t\} \in \Im} q_t(k_t) + \sum_{\{t,t'\} \in \Im} g_{t,t'}(k_t, k_{t'}). \tag{3}$$
   Verification of the subconvexity condition comes down to the verification of the subconvexity of every function $g_{t,t'}$ independently: $\forall r, r' \in L \setminus \{l\}$:
   $$g_{t,t'}(r+1, r'+1) + g_{t,t'}(r, r') \geq g_{t,t'}(r+1, r') + g_{t,t'}(r, r'+1).$$
2. The order of the structure is three and the number of labels is two [6].

The question of polynomial solvability for an arbitrary subconvex (max,+) problem remains open.

From the subconvexity condition (2) it follows that if k^*_T and k^{**}_T are solutions of a subconvex (max,+) problem then $k^*_T \bigvee k^{**}_T$ and $k^*_T \bigwedge k^{**}_T$ are also solutions of the same problem. Hence, one can define ***the highest*** and ***the lowest*** optimal labelings in the following way:

$$\breve{k}_T = \bigvee_{k^*_T = \arg\max\limits_{k_T} Q(k_T)} k^*_T\ , \qquad \frown{k}_T = \bigwedge_{k^*_T = \arg\max\limits_{k_T} Q(k_T)} k^*_T\ .$$

It can be shown that computational complexity of searching for the lowest as well as the highest optimal labelings is the same as for an arbitrary optimal labeling.

The following lemma describes a property of the subconvex (max,+) problems.

Lemma 1. *Let* $\overset{\frown}{k_T} = \bigwedge\limits_{k_T^* = \arg\max\limits_{k_T} Q(k_T)} k_T^*$ *be the lowest optimal labeling for some subconvex problem and* k_T *be an arbitrary labeling satisfying the condition:*

$$k_T \bigwedge \overset{\frown}{k_T} \neq \overset{\frown}{k_T} \tag{4}$$

Then the quality of the labeling k_T *is strictly less than the quality of the maximum of the labelings* k_T *and* $\overset{\frown}{k_T}$*:*

$$Q(k_T) < Q(k_T \bigvee \overset{\frown}{k_T}).$$

Proof. Let us rewrite the inequality (2) for labelings k_T and $\overset{\frown}{k_T}$

$$Q(k_T) + Q(\overset{\frown}{k_T}) \leq Q(k_T \bigwedge \overset{\frown}{k_T}) + Q(k_T \bigvee \overset{\frown}{k_T}) \tag{5}$$

The condition (4) together with the definition of the lowest optimal labeling leads to the inequality:

$$Q(k_T \bigwedge \overset{\frown}{k_T}) < Q(\overset{\frown}{k_T}). \tag{6}$$

We obtain the statement of the lemma by adding the inequalities (5) and (6).

4 The main result

We consider (max,+) problems with a structure of order two. In this case the quality function has the form (3), at that the functions q_t are arbitrary and the functions $g_{t,t'}(r, r') = \begin{cases} C_{t,t'} \geq 0, & r = r', \\ 0, & r \neq r'. \end{cases}$ This problem is not subconvex for more than two labels. Moreover, it is NP-hard (see [1]).

Let us build an auxiliary subconvex (max,+) problem for the given (max,+) problem. The exact solution of this auxiliary problem gives the optimal labels for the initial problem in some pixels.

This ***auxiliary problem*** is constructed in the following way:

1. The vision field T, the structure $\Im$ and the label set L are the same as in the initial problem.
2. Let us fix an arbitrary label $s \in L$.
3. Let us completely order labels in each pixel in the following way. The label s becomes the highest. Independently in each pixel one of the rest labels l_t^s whose quality $q_t(l_t^s)$ is maximal ($l_t^s = \arg\max\limits_{l \in L \setminus \{s\}} q_t(l)$) becomes the lowest.
4. The quality function of the auxiliary problem is

$$Q^s(k_T) = \sum_{\{t\} \in \Im} q_t(k_t) + \sum_{\{t,t'\} \in \Im} g^s_{t,t'}(k_t, k_{t'}),$$

where

$$g^s_{t,t'}(r,r') = \begin{cases} C_{t,t'},\ r = s, r' = s, \\ C_{t,t'},\ r \neq s, r' \neq s, \\ \ 0, \quad otherwise. \end{cases} \tag{7}$$

The constructed auxiliary problem is subconvex and its lowest solution $\widehat{k^s_T}$ can be found in polynomial time.

Theorem 1. *An arbitrary solution of the initial problem* $k^*_T = \arg\max\limits_{k_T} Q(k_T)$ *satisfies the following condition:* $k^*_T \bigwedge \widehat{k^s_T} = \widehat{k^s_T}$.

In other words the lowest optimal labeling for the initial problem in each pixel is not lower than the label in the corresponding pixel of the auxiliary problem solution ($\forall t \in T : k^*_t \wedge \widehat{k^s_t} = \widehat{k^s_t}$). The theorem allows to reduce the number of labels in each pixel. It is necessary to examine only those labels $l \in L$ in the pixel t that lie not lower than the label $\widehat{k^s_T}$: $L_t = \{l \in L : l \bigwedge \widehat{k^s_t} = \widehat{k^s_t}\}$. In particular, the set of remaining labels L_t can consist only of the highest label $\widehat{k^s_t} = s$. In this case the value of any optimal labeling of the initial problem in the pixel t equals s.

The algorithm based on the theorem consists in execution of the following three steps for every label $s \in L$:

1. Construction of the auxiliary problem for the label s;
2. Searching for the lowest solution $\widehat{k^s_T}$ of the auxiliary problem;
3. If $\widehat{k^s_t} = s$ then $k^*_t = s$.

Proof. We will conduct the proof of the theorem in the following way: we construct for any labeling k_T, such that $k_T \bigwedge \widehat{k^s_T} \neq \widehat{k^s_T}$ a labeling with better quality. Namely, $Q(k_T \bigvee \widehat{k^s_T}) > Q(k_T)$.

Let us consider an arbitrary labeling k_T, satisfying the condition $k_T \bigwedge \widehat{k^s_T} \neq \widehat{k^s_T}$. It follows from the lemma 1 that

$$0 < Q^s(k_T \bigvee \widehat{k^s_T}) - Q^s(k_T). \tag{8}$$

If $\widehat{k^s_t} \neq s$ then $\widehat{k^s_t} = l^s_t$. Otherwise one can construct a new labeling with the same quality, which is lower than $\widehat{k^s_T}$. Therefore, the lowest labeling $\widehat{k^s_T}$ of the auxiliary problem can take only two values in every pixel: $\widehat{k^s_t} \in \{s, l^s_t\}$.

The following inequality follows from formula (7) and aforesaid:

$$g^s_{t,t'}(k_t \vee \widehat{k^s_t}, k_{t'} \vee \widehat{k^s_{t'}}) - g^s_{t,t'}(k_t, k_{t'}) \leq g_{t,t'}(k_t \vee \widehat{k^s_t}, k_{t'} \vee \widehat{k^s_{t'}}) - g_{t,t'}(k_t, k_{t'}) \tag{9}$$

Using inequalities (8) and (9) we obtain:

$$0 < Q^s(k_T \bigvee \widehat{k^s_T}) - Q^s(k_T) =$$

$$= \sum_{t\in T} q_t(k_t \vee \widehat{k^s_t}) + \sum_{\{t,t'\}\in\Im} g^s_{t,t'}(k_t \vee \widehat{k^s_t}, k_{t'} \vee \widehat{k^s_{t'}}) - \left(\sum_{t\in T} q_t(k_t) + \sum_{\{t,t'\}\in\Im} g^s_{t,t'}(k_t, k_{t'}) \right) =$$

$$= \sum_{t\in T} \left(q_t(k_t \vee \widehat{k^s_t}) - q_t(k_t \vee \widehat{k^s_t}) \right) + \sum_{\{t,t'\}\in\Im} \left(g^s_{t,t'}(k_t \vee \widehat{k^s_t}, k_{t'} \vee \widehat{k^s_{t'}}) - g^s_{t,t'}(k_t, k_{t'}) \right) \leq$$

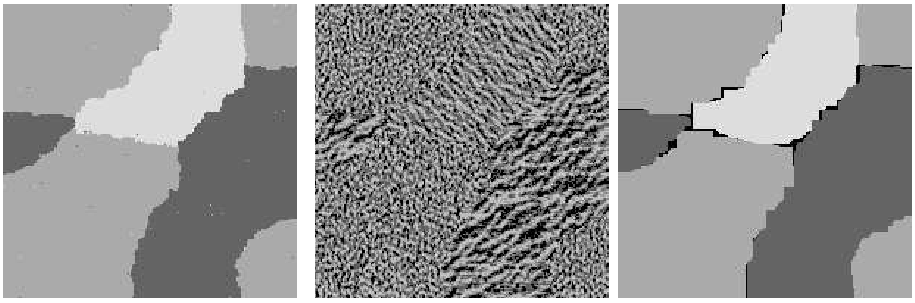
$$\leq \sum_{t\in T} \left(q_t(k_t \vee \widehat{k^s_t}) - q_t(k_t \vee \widehat{k^s_t}) \right) + \sum_{\{t,t'\}\in\Im} \left(g_{t,t'}(k_t \vee \widehat{k^s_t}, k_{t'} \vee \widehat{k^s_{t'}}) - g_{t,t'}(k_t, k_{t'}) \right) =$$

$$= Q(k_T \bigvee \widehat{k^s_T}) - Q(k_T).$$

Thus, $Q(k_T) < Q(k_T \bigvee \widehat{k^s_T})$ which was to be proved.

5 Experimental verification of the algorithm

The following three experiments demonstrate our approach for segmentation and stereoreconstruction problems. In all cases the vision field is a rectangular area of the integer-valued lattice $T = \{(i,j) | 0 \leq i < I, 0 \leq j < J\}$.



(a) Ground truth (b) Generated artificial image (c) 98.8% of the optimal labeling was found

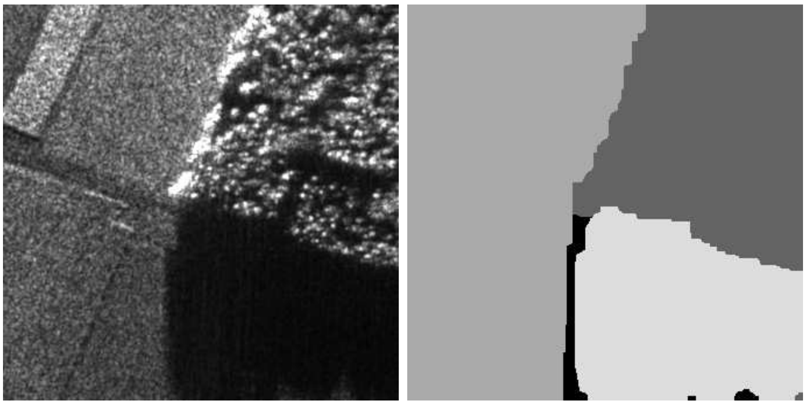
Fig. 1. Segmentation of the artificial image

The structure is $\Im = \{\{(i,j),(i+1,j)\} \subset T, \{(i,j),(i,j+1)\} \subset T\}$. The quality function has the form (3) with the functions

$$g_{t,t'}(r,r') = \begin{cases} C \geq 0 & r = r', \\ 0 & r \neq r'. \end{cases}$$

This problem is NP-hard if the number of labels is greater than two (see [1]).

The first two examples (see fig. 1 and 2) demonstrate applicability of the algorithm for texture segmentation. In the first example an image with three textures was generated using a Markov Random Field model (fig. 1(b)). Due to this fact the real segmentation (fig. 1(a)) is known. The second example (fig. 2) demonstrates texture segmentation of a real image (fig. 2(a)). The number $q_t(k)$ defines a probability that pixel t belongs to the k-th ($k \in \{1, 2, 3\}$) texture. Figures 1(c) and 2(b) present the results of the optimal labeling search. Black color is used to mark those pixels where answer "no label" was obtained.



(a) Initial image

(b) 98.4% of the optimal labeling was found

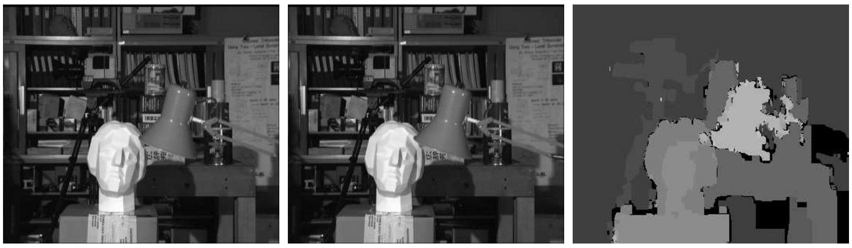
Fig. 2. Segmentation of the airphoto

The third example (see fig. 3) demonstrates the problem of 3D reconstruction by stereoimagepairs (fig. 3(a,b)). The qualities $q_t(k)$ are found in some reasonable way for all pixels t and 20 possible disparities ($k \in \{1, ..., 20\}$). Then the 3D reconstruction problem is formulated as a (max,+) problem. The part of the optimal labeling found by our approach is shown in the figure 3(c). Again black color is used to mark those pixels where answer "no label" was obtained.

Of course, the quality of reconstruction can be improved using a more thorough choice of the numbers $q_t(k)$, however, it is not a goal of our research.

6 Conclusion

We have shown in this paper that the exact solution can be found at least partially even for NP-hard (max,+) problems. Experimental results show the possibility to restore the optimal labeling for almost all pixels.



(a) Left picture

(b) Right picture

(c) 93.6% of the optimal labeling was found

Fig. 3. Reconstruction of three-dimensional scene by the pair of stereoimages

References

1. Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
2. S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. on PAMI*, 6(6):721–741, 1984.
3. D. M. Greig, B. T. Porteous, and A. H. Seheult. Exact maximum a posteriori estimation for binary images. *J. Royal Statistical Soc., Series B*, 51(2):271–279, 1989.
4. Hiroshi Ishikawa and Davi Geiger. Segmentation by grouping junctions. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998.
5. S. Kirkpatrick, C. D. Gellatt, Jr., and M. P. Vecch. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.
6. Vladimir Kolmogorov and Ramin Zabih. What energy functions can be minimized via graph cuts. In A. Heyden et al., editor, *ECCV 2002*, number 2352 in LNCS, pages 65–81, Berlin Heidelberg, 2002. Springer-Verlag.
7. Stan Z. Li. *Markov Random Field Modeling in Image Analysis*. Computer Science, Workbench, Springer, 2001.
8. M. I. Schlesinger. Syntax analysis of two dimensional visual signals with noise. *Cybernetics, Kiev*, 4:113–130, 1976. in russian.
9. M. I. Schlesinger and B. Flach. Some solvable subclass of structural recognition problems. In Tomas Svoboda, editor, *Czech Pattern Recognition Workshop 2000*, pages 55–61, Praha, February 2000. Czech Pattern Recognition Society.
10. M. I. Schlesinger and B. Flach. Analysis of optimal labelling problems and their applications to image segmentation and binocular stereovision. In Franz Leberl and Andrej Ferko, editors, *Proceedings East-West-Vision 2002 (EWV'02)*, pages 55–60. International Workshop and Project Festival on Computer Vision, Computer Graphics, New Media, 2002.
11. M. I. Schlesinger and V. K. Koval. Two dimensional programming in image analysis problems. *Automatics and Telemechanics, Moscow*, 2:149–168, 1976. in russian.