6. Peculiarity of power minimisation at labelling of two-dimensional images

The results described in the previous sections can be summarised with following two theorems, which are merely reformulation of the lemma 3 and theorem 4 in a slightly different form. This new formulation is given only for convenience of the following narrative.

Theorem 5. Let $N(t)=\{t'\in T \mid tt'\in\Im\}$ be a set of objects neighbouring with t and Φ be an array $\left(\varphi_{tt'}(k) \mid t\in T,\ t'\in N(t),\ k\in K\right)$ of potentials.

Let z and z' be two $(\max,+)$-problems defined for sets T, K and $\Im$, first of them with weights $g_{tt'}(k,k')$ and $q_t(k)$ and second with weights $g'_{tt'}(k,k')$ and $q'_t(k)$.

Let a graph formed by a neighbourhood $\Im$ on the set T be connected.

The problems z and z' are equivalent if and only if an array Φ of potentials exists such that

$$g'_{tt'}(k,k')=g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k'),\ tt'\in\Im,\ k\in K,\ k'\in K,$$

$$q'_t(k)=q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k),\ t\in T,\ k\in K.$$

■

Theorem 6. Let z be a $(\max,+)$-problem defined for sets T, K and $\Im$ with weights $g_{tt'}(k,k')$ and $q_t(k)$. Let the problem z have a trivial equivalent. Let Φ' be an array of potentials $\varphi'_{tt'}(k)$, $tt'\in\Im$, $k\in K$, that minimises a problem power

$$E(\Phi)=\sum_{tt'\in\Im}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+\sum_{t\in T}\max_{k}\left(q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k)\right),$$

$$E(\Phi')=\min_{\Phi}E(\Phi).$$

In this case the problem z' defined for the sets T, K, $\Im$ with weights

$$g'_{tt'}(k,k')=g_{tt'}(k,k')+\varphi'_{tt'}(k)+\varphi'_{t't}(k'),\ tt'\in\Im,\ k\in K,\ k'\in K,$$

$$q'_t(k)=q_t(k)-\sum_{t'\in N(t)}\varphi'_{tt'}(k),\ t\in T,\ k\in K,$$

is trivial.

■

In the previous sections the general algorithm is described for power minimisation at arbitrary $(\max,+)$-problem. Now the power minimisation will be analysed at a special case of neighbourhood $\Im$, which is specific for visual images. The special case of $(\max,+)$-problems will be considered when the set T of objects is a set of pixels, i.e. rectangular region of two-dimensional integer grid. Namely,

$$T=\left\{(i,j)\middle|1\le i\le m,\ 1\le j\le n\right\},$$

where m is a number of rows and n is a number of columns of an image. It will be assumed that the neighbourhood $\Im$ is natural for visual images set

$$\left\{\{(i-1,j),(i,j)\}\middle|1<i\le m,\ 1\le j\le n\right\}\cup\left\{\{(i,j-1),(i,j)\}\middle|1\le i\le m,\ 1<j\le n\right\}$$

of pixel pairs. As it can be seen, the set $\Im$ consists of two non-intersecting subsets: subset

$$\Im^V=\left\{\{(i-1,j),(i,j)\}\middle|1<i\le m,\ 1\le j\le n\right\}$$

of vertically neighbouring pixel pairs and subset

$$\Im^H=\left\{\{(i,j-1),(i,j)\}\middle|1\le i\le m,\ 1<j\le n\right\}$$

of horizontally neighbouring pairs,

$$\Im=\Im^H\cup\Im^V.$$

Similarly, the set $N(t)$ of neighbours of the pixel t consists of two non-intersecting subsets: subset $N^H(t)$ of horizontal neighbours and subset $N^V(t)$ of vertical neighbours. It has to be mentioned that neither $\Im^H$ nor $\Im^V$ contains cycle and so calculation of the values

$$\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^H}g_{tt'}\left(k(t),k(t')\right)+\sum_{t\in T}q_t\left(k(t)\right)\right],$$

$$\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^V}g_{tt'}\left(k(t),k(t')\right)+\sum_{t\in T}q_t\left(k(t)\right)\right]$$

can be easily fulfilled with dynamic programming.

Due to the mentioned peculiarities, the minimisation of the problem power

$$E(\Phi)=\sum_{tt'\in\Im}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+\sum_{t\in T}\max_{k}\left(q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k)\right)$$

can be fulfilled more effectively than in general case of an arbitrary $(\max,+)$-problem.

Let us consider an array Ψ of numbers $\psi_t(k)$, $t\in T$, $k\in K$, which will be also referred to as potentials.

Theorem 7. It holds the equality

$$\min_{\Phi}\left[\sum_{tt'\in\Im}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+\sum_{t\in T}\max_{k}\left(q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k)\right)\right]=$$

$$=\min_{\Psi}\left[\max_{\bar{k}\in K^T}\left(\sum_{tt'\in\Im^H}g_{tt'}(k(t),k(t'))+\frac{1}{2}\sum_{t\in T}q_t(k(t))+\sum_{t\in T}\psi_t(k(t))\right)+\right.$$

$$\left.+\max_{\bar{k}\in K^T}\left(\sum_{tt'\in\Im^V}g_{tt'}(k(t),k(t'))+\frac{1}{2}\sum_{t\in T}q_t(k(t))-\sum_{t\in T}\psi_t(k(t))\right)\right]. \qquad (29)$$

■

Proof. 1. Let us consider the second sum in the left part of (29). For this sum the following equality and inequality are valid:

$$\sum_{t\in T}\max_{k}\left(q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k)\right)=$$

$$=\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)+\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right)\le$$

$$\le\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)\right)+$$

$$+\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right). \qquad (30)$$

2. The inequality (30) holds for arbitrary values of potentials $\psi_t(k)$. However, for some of their values the inequality becomes an equality. It occurs when

$$\psi_t(k)=\frac{1}{2}\left[\sum_{t'\in N^H(t)}\varphi_{tt'}(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right].$$

So, it holds an equality

$$\sum_{t\in T}\max_{k}\left(q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k)\right)=$$

$$=\min_{\Psi}\left[\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)\right)+\right.$$

$$\left.+\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right)\right]. \qquad (31)$$

3. Let us designate Q the left part of (29) and write it taking into account the proved equality (31):

$$Q=\min_{\Phi}\left\{\sum_{tt'\in\Im}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+\right.$$

$$+\min_{\Psi}\left[\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)\right)+\right.$$

$$\left.\left.+\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right)\right]\right\}. \qquad (32)$$

4. An array

$$\Phi=\left(\varphi_{tt'}(k)\ \middle|\ t\in T,\ t'\in N(t),\ k\in K\right)$$

can be regarded as a pair $\left(\Phi^H,\Phi^V\right)$ of two arrays

$$\Phi^H=\left(\varphi_{tt'}(k)\ \middle|\ t\in T,\ t'\in N^H(t),\ k\in K\right),$$

$$\Phi^V=\left(\varphi_{tt'}(k)\ \middle|\ t\in T,\ t'\in N^V(t),\ k\in K\right).$$

Then the equality (32) can be transformed using evident rules

$$\min_{\Phi}\left[f_1(\Phi)+\min_{\Psi}f_2(\Phi,\Psi)\right]=\min_{\Psi}\min_{\Phi}\left[f_1(\Phi)+f_2(\Phi,\Psi)\right],$$

$$\min_{\Phi}f(\Phi)=\min_{\Phi^H}\min_{\Phi^V}f\left(\Phi^H,\Phi^V\right),$$

$$\sum_{tt'\in\Im}f_{tt'}=\sum_{tt'\in\Im^H}f_{tt'}+\sum_{tt'\in\Im^V}f_{tt'}.$$

$$Q=\min_{\Psi}\min_{\Phi^H}\min_{\Phi^V}\left\{\sum_{tt'\in\Im^H}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+\right.$$

$$+\sum_{t\in T}\max_k\left[\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)\right)\right]+$$

$$+\sum_{tt'\in\Im^V}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+$$

$$+\sum_{t\in T}\max_k\left[\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right]\Bigg\}.$$

5. The last expression can be transformed using an evident rule

$$\min_{\Phi^H}\min_{\Phi^V}\left(f_1\left(\Phi^H\right)+f_2\left(\Phi^V\right)\right)=\min_{\Phi^H}f_1\left(\Phi^H\right)+\min_{\Phi^V}f_2\left(\Phi^V\right).$$

$$Q=\min_{\Psi}\Bigg\{\min_{\Phi^H}\Bigg[\sum_{tt'\in\Im^H}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+$$

$$+\sum_{t\in T}\max_k\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)\right)\Bigg]+$$

$$+\min_{\Phi^V}\Bigg[\sum_{tt'\in\Im^V}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+$$

$$+\sum_{t\in T}\max_k\left(\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right)\Bigg]\Bigg\}. \qquad (33)$$

6. The first term

$$\min_{\Phi^H}\Bigg[\sum_{tt'\in\Im^H}\max_{k,k'}\left(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right)+$$

$$+\sum_{t\in T}\max_k\left(\frac{1}{2}q_t(k)+\psi_t(k)-\sum_{t'\in N^H(t)}\varphi_{tt'}(k)\right)\Bigg] \qquad (34)$$

in (33) relates to $(\max,+)$-problem z^* defined for the set T, acyclic neighbourhood $\Im^H$ and weights

$$g_{tt'}(k,k'),\ tt'\in\Im^H,\ k\in K,\ k'\in K, \text{ and } \frac{1}{2}q_t(k)+\psi_t(k),\ t\in T,\ k\in K.$$

The expression (34) means a minimal power on the set of problems, which are equivalent to z^*. The neighbourhood $\Im^H$ contains no cycle and so, due to the theorem 1, the problem z^* is equivalent to trivial one and searched minimum (34) equals the quality

$$\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^H} g_{tt'}\big(k(t),k(t')\big)+\frac{1}{2}\sum_{t\in T} q_t\big(k(t)\big)+\sum_{t\in T}\psi_t\big(k(t)\big)\right]$$

of the best labelling in the $(\max,+)$-problem with neighbourhood $\Im^H$. Similarly, the second term

$$\min_{\Phi^V}\left[\sum_{tt'\in\Im^V}\max_{k,k'}\big(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\big)+\right.$$
$$\left.+\sum_{t\in T}\max_{k}\left(\frac{1}{2}q_t(k)-\psi_t(k)-\sum_{t'\in N^V(t)}\varphi_{tt'}(k)\right)\right]$$

in (33) equals the quality

$$\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^V} g_{tt'}\big(k(t),k(t')\big)+\frac{1}{2}\sum_{t\in T} q_t\big(k(t)\big)-\sum_{t\in T}\psi_t\big(k(t)\big)\right]$$

of the best labelling in other $(\max,+)$-problem, namely, in the problem with neighbourhood $\Im^V$.

So,

$$Q=\min_{\Psi}\left\{\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^H} g_{tt'}\big(k(t),k(t')\big)+\frac{1}{2}\sum_{t\in T} q_t\big(k(t)\big)+\sum_{t\in T}\psi_t\big(k(t)\big)\right]+\right.$$
$$\left.+\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^V} g_{tt'}\big(k(t),k(t')\big)+\frac{1}{2}\sum_{t\in T} q_t\big(k(t)\big)-\sum_{t\in T}\psi_t\big(k(t)\big)\right]\right\}.$$

■

7. Algorithm for power minimisation.

Due to the theorem 7 minimisation of the problem power

$$E(\Phi)=\sum_{tt'\in\Im}\max_{k,k'}\big(g_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\big)+\sum_{t\in T}\max_{k}\left(q_t(k)-\sum_{t'\in N(t)}\varphi_{tt'}(k)\right)$$

with respect to potentials $\varphi_{tt'}(k)$, $t\in T$, $t'\in N(t)$, $k\in K$, is reduced to minimisation of the power

$$Q(\Psi)=\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^H} g_{tt'}\big(k(t),k(t')\big)+\frac{1}{2}\sum_{t\in T}\Big(q_t\big(k(t)\big)+\psi_t\big(k(t)\big)\Big)\right]+$$
$$+\max_{\bar{k}\in K^T}\left[\sum_{tt'\in\Im^V} g_{tt'}\big(k(t),k(t')\big)+\frac{1}{2}\sum_{t\in T}\Big(q_t\big(k(t)\big)-\psi_t\big(k(t)\big)\Big)\right] \qquad (35)$$

with respect to potentials $\psi_t(k)$, $t\in T$, $k\in K$.

For each labelling $\bar{k}:T\to K$ two qualities are defined: horizontal quality

$$G^H\left(\bar{k},\Psi\right)=\sum_{tt'\in\Im^H} g_{tt'}\left(k(t),k(t')\right)+\frac{1}{2}\sum_{t\in T}\left(q_t\left(k(t)\right)+\psi_t\left(k(t)\right)\right)$$

and vertical quality

$$G^V\left(\bar{k},\Psi\right)=\sum_{tt'\in\Im^V} g_{tt'}\left(k(t),k(t')\right)+\frac{1}{2}\sum_{t\in T}\left(q_t\left(k(t)\right)-\psi_t\left(k(t)\right)\right).$$

The horizontal quality $G^H\left(\bar{k},\Psi\right)$ of each labelling $\bar{k}$ is a linear and, consequently, a convex function of potentials Ψ, as well as its vertical quality $G^V\left(\bar{k},\Psi\right)$. So, the functions

$$G^H\left(\Psi\right)=\max_{\bar{k}\in K^T} G^H\left(\bar{k},\Psi\right),\quad G^V\left(\Psi\right)=\max_{\bar{k}\in K^T} G^V\left(\bar{k},\Psi\right)$$

are also convex functions of potentials Ψ. Finally, the power

$$Q\left(\Psi\right)=\max_{\bar{k}\in K^T} G^H\left(\bar{k},\Psi\right)+\max_{\bar{k}\in K^T} G^V\left(\bar{k},\Psi\right) \tag{36}$$

is also convex function of potentials as a sum of two convex functions. It is important to mention that each terms in expressions (35) and (36) can be easily calculated, as well as the labellings

$$\bar{k}^{*H}=\arg\max_{\bar{k}\in K^T} G^H\left(\bar{k},\Psi\right),\quad \bar{k}^{*V}=\arg\max_{\bar{k}\in K^T} G^V\left(\bar{k},\Psi\right)$$

because all calculations mean the solution to two acyclic $\left(\max,+\right)$-problems.

Due to convexity of the power $Q\left(\Psi\right)$ it can be minimised with the method of sub-gradient descent just as it has been done for power minimisation for $\left(\max,+\right)$-problem of general form. For the problems under present consideration the sub-gradient descent consists in the following.

Let γ_i, $i=0,1,...,$ be a sequence of numbers such that

$$\lim_{i\to\infty}\gamma_i=0,\ \lim_{n\to\infty}\sum_{i=0}^{n}\gamma_i=\infty.$$

Let Ψ_0 be an arbitrary array of potentials $\psi_t\left(k\right)$, $t\in T$, $k\in K$, for example,

$$\psi_t\left(k\right)=0.$$

Let Ψ_i, $i=1,2,...,$ be a sequence of arrays such that

$$\Psi_{i+1}=\Psi_i-\gamma_i\cdot\nabla\Psi_i,$$

where $\nabla\Psi_i$ is a sub-gradient of the function $Q\left(\Psi\right)$ at point Ψ_i.

In that case

$$\lim_{i\to\infty}Q\left(\Psi_i\right)=\min_{\Psi}Q\left(\Psi\right).$$

To apply this method for our case it is necessary only to define how a sub-gradient of the function $Q(\Psi)$ has to be calculated at arbitrary given array Ψ. These calculations are quite simple. A sub-gradient $\nabla\Psi$ is an array of numbers $\nabla\psi_t(k)$, $t \in T$, $k \in K$, and each of these numbers has to be calculated in a following way.

Let

$$\bar{k}^H = \arg\max_{\bar{k}\in K^T} G^H(\bar{k},\Psi),\ \bar{k}^V = \arg\max_{\bar{k}\in K^T} G^V(\bar{k},\Psi).$$

When the labelling $\bar{k}^H$ is not defined unambiguously, i.e. if there are several best labellings, any of them can be chosen as the labelling $\bar{k}^H$. The same holds for the labelling $\bar{k}^V$. For the chosen labellings $\bar{k}^H$ and $\bar{k}^V$ the components $\nabla\psi_t(k)$ of the sub-gradients are assigned the values

$$\nabla\psi_t(k) = 1, \text{if } k^H(t) = k,\ k^V(t) \neq k,$$

$$\nabla\psi_t(k) = -1, \text{if } k^H(t) \neq k,\ k^V(t) = k,$$

$$\nabla\psi_t(k) = 0 \quad \text{otherwise.}$$

8. Illustrations of proposed approach

8.1. More precise definition of proposed algorithms

Two algorithms for solution to $(\max,+)$-problems via their equivalent transformation have been described. The first of them, let us call it Algorithm 1, minimises a power of arbitrary $(\max,+)$-problem. The second one, just now described, let us call it Algorithm 2, minimises a power of $(\max,+)$-problems, whose object set T is a rectangular region of two-dimensional integer grid. Both algorithms are defined up to a sequence γ_i, $i = 1,2,...,$, of step values. Moreover, both algorithms are not defined completely because no stop condition was defined for them.

All below described examples, with exception of the example in the section 4.2, are obtained with an harmonic series γ_i, $i = 1,2,...$, $\gamma_i = 1/i$. The algorithms were stopped when a zero value of sub-gradient was achieved, the gradient being calculated as it was said in the sections 4 and 7. We are far from certainty that just such algorithms specification is the best one in a sense. Just such specification was made only because for running the program the steps values and stop condition should be at least somehow specified. As to optimisation of steps γ_i and stop condition, it will require further and obviously tiresome researches.

8.2. Solution to $(\vee, \&)$-problems via power minimisation

It was mentioned in the Section 5 that the proposed power minimisation approach shows a way to manage not only $(\max, +)$-problems but $(\vee, \&)$-problems too, this approach being stronger than widely used for this aim relaxation labelling [23]. Let us show an example that illustrates this statement. The following $(\vee, \&)$-problem presented on the fig.1 will be considered.

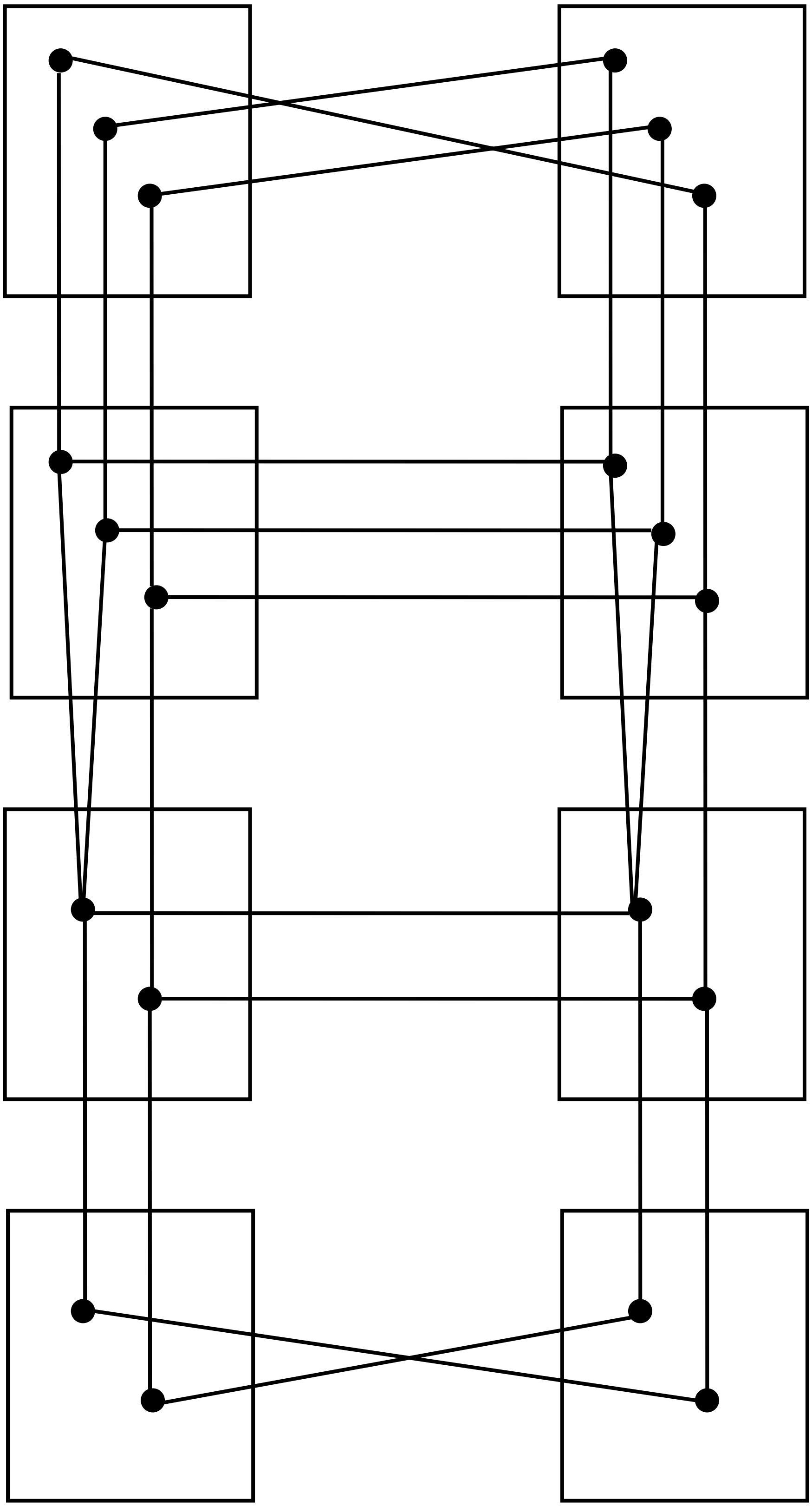


Fig.1

On the figure the set T is presented with squares. The set $\mathfrak{I}$ of neighbouring object pairs includes the pairs of squares, which are neighbours in a horizontal direction, as well as vertically neighbouring pairs. So, the set $\Im$ consists of 8 pairs: 4 pairs of horizontal neighbours and 6 pairs of vertical ones. The set K of labels is not explicitly presented on the figure, there are presented only numbers $q_t(k)$, $t \in T$, $k \in K$, which in $(\vee,\&)$-problem take values 0 or 1. Namely, there are presented with black dots only those labels $k \in K$, for which $q_t(k)=1$. So, it is seen from the figure that for some objects 3 labels are allowable and only 2 for others. The numbers $g_{tt'}(k,k')$ are represented with edges joining some dots in neighbouring squares.

In terms of this pictorial presentation of an $(\vee,\&)$-problem a labelling is meant as a choice of one and only one dot in each square, the labelling being allowable if the dots chosen in neighbouring squares are joined with an edge. The $(\vee,\&)$-problem presented on the figure consists in answering the question whether such choice is possible for the given figure or not.

A widely used solution to this problem is based on necessary, not sufficient, condition of positive solution to $(\vee,\&)$-problem, which was mentioned in the Section 5. Namely, for positive solution to $(\vee,\&)$-problem it is necessary that it contains a non-empty locally consistent part. This condition can be easily examined with a famous relaxation labelling. It is easy to see that this condition is satisfied for presented example because the given problem itself is non-empty and locally consistent. It is also easy to see that the presented problem has no positive solution that shows insufficiency of the quoted condition. An idea of $(\max,+)$-problem power minimisation allows to test the positive solvability of $(\vee,\&)$-problem more strongly than with quoted condition. This test consists in the following.

For a given $(\vee,\&)$-problem with numbers $g_{tt'}(k,k') \in \{0,1\}$ and $q_t(k) \in \{0,1\}$ an auxiliary $(\max,+)$-problem has to be considered with such weights $\overline{g}_{tt'}(k,k') \in R$ and $\overline{q}_t(k) \in R$ that

$$\overline{g}_{tt'}(k,k') = 0 \text{ if } g_{tt'}(k,k') = 1,$$

$$\overline{g}_{tt'}(k,k') = -L \text{ if } g_{tt'}(k,k') = 0,$$

$$\overline{q}_t(k) = 0 \text{ if } q_t(k) = 1,$$

$$\overline{q}_t(k) = -L \text{ if } q_t(k) = 0,$$

where L is an arbitrary positive number, for example, 100. Then it has to be tested whether a negative power can be achieved on the set of $(\max,+)$-problem, which are equivalent to given auxiliary problem. If yes then there is no labelling with zero quality in auxiliary $(\max,+)$-problem and, consequently, the given $(\vee,\&)$-has no positive solution.

An implementation of this approach for the problem under consideration is presented on the fig. 2. The certain array of potentials $\varphi_{tt'}(k)$ is shown on the picture with little triangles near the dots, which represent labels. The potentials take values -3, -2, -1, 0, +1, +2, +3. These numbers are represented with one triangle or with a pair or with a triple of triangles. An absence of triangles near some dot means a zero value of the potential. The triangles have two possible orientations: with a vertex of triangle toward and outward a corresponding dot. At first case the triangles express a negative value of potentials, which increases a quality of a dot. Triangles, which are oriented with their vertex outward the dot, mean a positive value of potentials, which decreases a dot's quality.

The potentials on fig. 2 transform the auxiliary $(\max,+)$-problem into equivalent one but with power (-1). Really, for each pairs tt' of neighbouring squares a value $\max\limits_{k,k'}\left[\overline{g}_{tt'}(k,k')+\varphi_{tt'}(k)+\varphi_{t't}(k')\right]$ remains to be zero. For the square t^* in a second from above row and a left column the value $\max\limits_{k}\left[\overline{q}_{t^*}(k)-\sum\limits_{t'\in N(t^*)}\varphi_{tt^*}(k)\right]$ becomes (-1). For all other squares $t\neq t^*$ a value $\max\limits_{k}\left[\overline{q}_{t}(k)-\sum\limits_{t'\in N(t)}\varphi_{tt'}(k)\right]$ remains zero.

The potentials presented on the fig. 2 were found manually. Of course, they can be found automatically both with the Algorithm 1 and with the Algorithm 2. The Algorithm 1 was used and during 11 iterations it transformed the auxiliary $(\max,+)$-problem into equivalent one with negative power and so an absence of positive solution to initial $(\vee,\&)$-problem was stated.

The considered approach can be slightly complicated to emphasise the main difference between proposed approach and known ones.

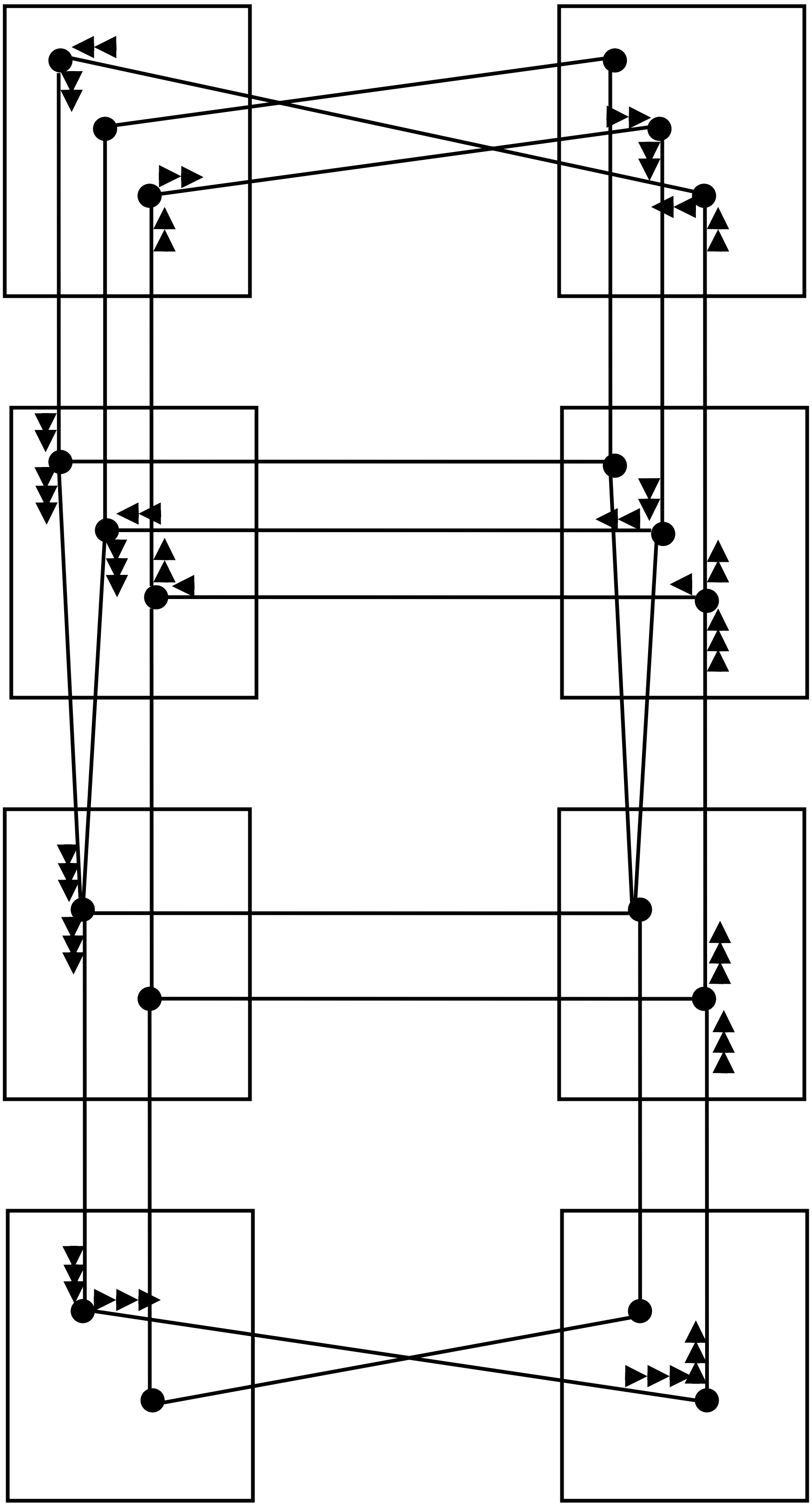


Fig. 2

Let us consider the $\left(\max,+\right)$-problem, presented on a fig. 3.

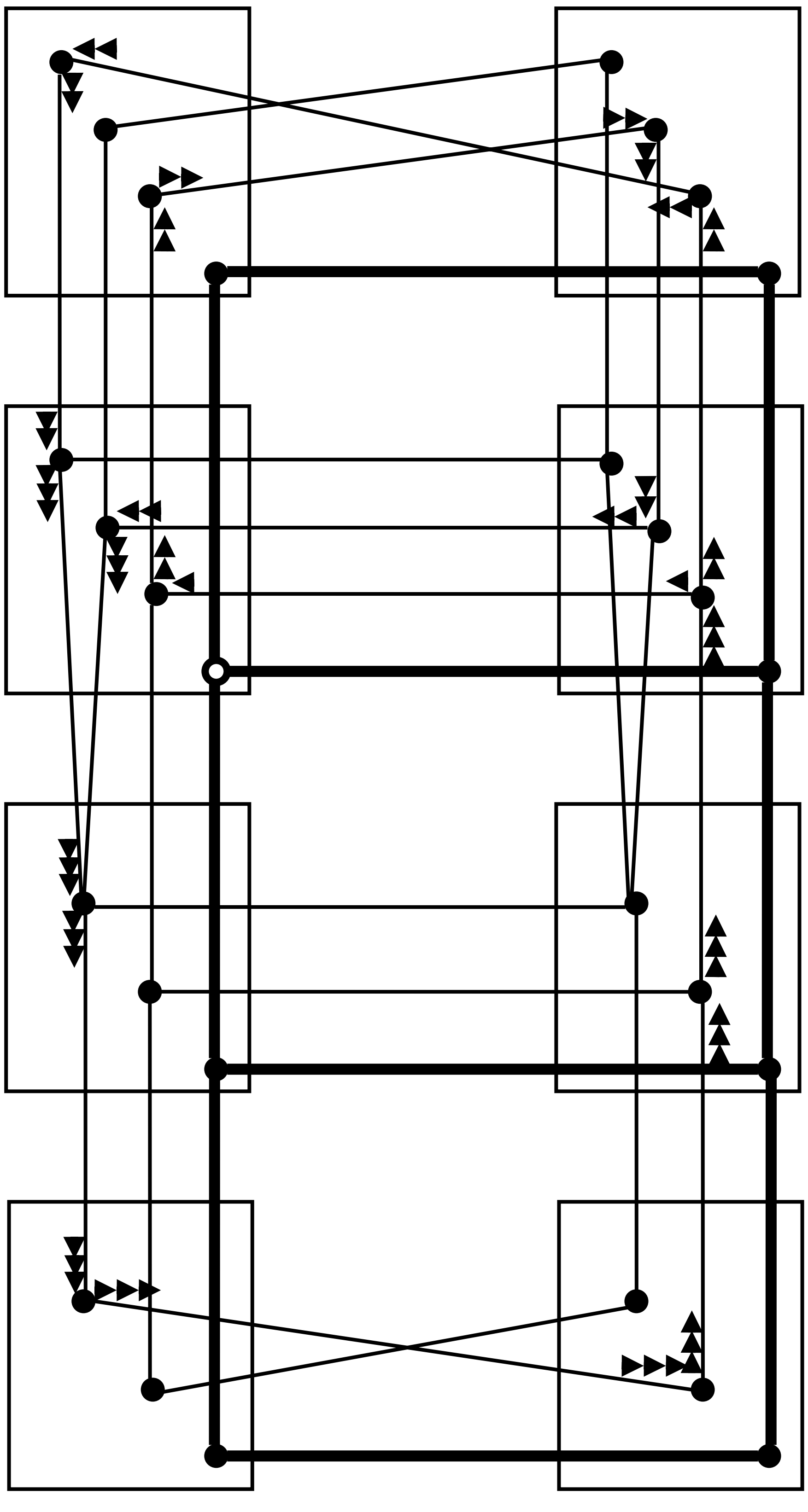


Fig. 3

A set T is presented here with 8 squares just as in previous example and with the same neighbourhood $\Im$. In each square t some dots are drawn, which represent various pairs (t,k) object-label. Let t^* be the second from above square in the left column, k^* be a label, designated with a white dot and let $q_{t^*}(k^*) = -1$. Let the values $q_t(k)$ equal 0 for all other pairs (t,k) as well as values $g_{tt'}(k,k')$ for all quadruples $t,\ t',\ k,\ k'$, designated on the figure with edges. For all others quadruples $t,\ t',\ k,\ k'$ the value $g_{tt'}(k,k')$ equals some big negative number, let us say (-100).

The defined $(\max,+)$-problem is not trivial but is equivalent to trivial one. The potentials, which transform the problem into trivial one, are depicted on a fig. 3. At these potentials a power of transformed problem equals (-1) and this value is just a quality of the best labelling, represented on a fig. 3 with thick lines. Such or other potentials, which transform the initial problem into trivial one, will be obligatorily found with an algorithm, which minimises a problem power, while none of the known algorithms is able to provide such trivialisation. Really, the initial $(\max,+)$-problem is such that the corresponding $(\vee,\&)$-problem already contains a non-empty locally consistent part. The known up to now algorithms stop at this condition and, consequently, can find no trivial equivalent to the problem under solution even if such equivalent exists. For algorithms minimising problem power such omission is impossible.

For trivialisation of presented problem the Algorithm 1 was used. The time, which was necessary for trivialisation, depended essentially on steps γ_i. Depending on them, required time was in a range from hundreds of iterations to very large number of them. Rather decent times were obtained at $\gamma_i = 100/i$ and even at constant γ_i, equal, for example, 0,2. One can see that the restrictions $\lim\limits_{i\to\infty}\gamma_i = 0$ and $\lim\limits_{n\to\infty}\sum\limits_{i=0}^{n}\gamma_i = \infty$ on steps γ_i, which follow from the general theory of sub-gradient descent, are rather weak. These restrictions embrace rather descent solutions, it is true, but completely unacceptable decisions too. As it was already mentioned above, the question about optimal sequence of steps γ_i remains opened.

8.3. Solution to $(\max,+)$-problems with trivial equivalent

A domain of applicability of proposed approach is a subclass of $(\max,+)$-problems, which have a trivial equivalent. It is worthy to emphasise once more that this class is rather wide since it contains all acyclic and all supermodular $(\max,+)$-problems. First of all, the developed algorithms were tested on the $(\max,+)$-problems of the domain of their competence. The aim of the experiments consisted not in testing an ability of algorithms to trivialise such problem because this fact was proved theoretically. The aim of experiments was to estimate a speed of algorithms, to compare the Algorithms 1 and 2, i.e. to answer the question, which were not theoretically researched, particularly, to check up how the assumed stop condition behaves that also was not theoretically deduced.

Results of experiments are summarised in the Table 1. Let us explain the data in the table by the example of fifth row of the table.

Table 1. Algorithms performance comparison

Problem size			Algorithm 1		Algorithm 2	
rows	columns	labels	number of iterations	time (sec)	number of iterations	time (sec)
5	5	2	194,5	1	81,7	<< 1
5	5	5	415,3	6	180,0	<< 1
5	5	10	506,1	26	192,4	<< 1
10	10	2	251,3	13	155,8	<< 1
10	10	5	421,7	59	198,6	<< 1
10	10	10	534,0	199	201,6	<< 1
50	50	2	1243,0	876	160,0	1
50	50	5	-	-	202,3	2
50	50	10	-	-	204,5	7
100	100	10	-	-	205,2	33

A set T of objects is a set of pixels that consists of 10 rows and 10 columns. A set $\Im$ is a conventional for pictures 4-neighbourhood. A set K of labels consists of 5 labels. At each experiment some labelling $\overline{k}^*$ was chosen beforehand and then the numbers $q_t(k)$ and

$g_{tt'}(k,k')$ were defined in such way that only the chosen labelling $\overline{k}^*$ became optimal. These numbers were defined so that

$$q_t(k) = 0 \text{ for each } t \in T \text{ and } k \in K,$$
$$g_{tt'}(k,k') = 0 \text{ if } k = k^*(t),\ k' = k^*(t'),$$
$$g_{tt'}(k,k') = -1 \text{ otherwise.}$$

It is evident that the problem defined in such way is trivial. This problem was equivalently transformed with random potentials $\varphi_{tt'}(k)$ uniformly distributed at the range from -50 to +50. So, the transformed problem was defined with weights

$$g'_{tt'}(k,k') = g_{tt'}(k,k') + \varphi_{tt'}(k) + \varphi_{t't}(k'),$$
$$q'_t(k) = -\sum_{t' \in N(t^*)} \varphi_{tt'}(k),$$

and as a rule became non-trivial. The obtained non-trivial problem was solved with the Algorithms 1 and 2. Each of them trivialised the problem and stopped after such or other number of iterations. For the given image size 10 by 10 this experiment was performed 20 times at different values of randomly chosen potentials $\varphi_{tt'}(k)$. It took in average 59 seconds and 421,7 iterations to trivialise a problem with Algorithm 1. An average duration of a trivialisation with Algorithm 2 was so small that it could not be measured. This fact is depicted in corresponding item of the table with an entry <<1. An average number of iteration equalled 198,6.

All other lines in the table have a similar sense. They were performed experiments for vision fields in a range from 5x5 to 100x100 and for 2, 5 and 10 labels. As it is seen from the first seven lines of the table the Algorithm 2 has an evident advantage as compared with the Algorithm 1 and so, for last three lines, the Algorithm 1 was not tested at all.

8.4. Recognition of one class of images

Let us define a set of images, which will be referred to as ideal images. Let T be a rectangular vision field with m rows and n columns, $T = \{(i,j) | 1 \le i \le m,\ 1 \le j \le n\}$. Let us define $m+n$ images of the form $T \to \{0,1\}$, which will be denoted as $e^V_{j^*}$, $j^* = 1,\ 2,\ \ldots,\ n$, and $e^H_{i^*}$, $i^* = 1,\ 2,\ \ldots,\ m$, so that

$$e^V_{j^*}(i,j)=1 \text{ if } j=j^*,\ i=1,\ 2,\ \dots,\ m,$$

$$e^V_{j^*}(i,j)=0 \text{ if } j\neq j^*,\ i=1,\ 2,\ \dots,\ m,$$

$$e^H_{i^*}(i,j)=1 \text{ if } i=i^*, j=1,\ 2,\ \dots,\ n,$$

$$e^H_{i^*}(i,j)=0 \text{ if } i\neq i^*, j=1,\ 2,\ \dots,\ n.$$

In other words, $e^V_{j^*}$ is an image of a vertical line where each pixel of j^*-th column is black and all other pixels are white. Similarly, $e^H_{i^*}$ is an image of an horizontal line on i^*-th row. Let us define also a completely white image e_0 with all pixels white. Let a disjunction of two images $e_1:T\rightarrow\{0,1\}$ and $e_2:T\rightarrow\{0,1\}$ be defined as such image $e:T\rightarrow\{0,1\}$ that

$$e(i,j)=e_1(i,j)\vee e_2(i,j).$$

A set E of ideal images is defined as a set that contains the completely white image e_0, the images $e^V_{j^*}$, $j^*=1,\ 2,\ \dots,\ n$, and $e^H_{i^*}$, $i^*=1,\ 2,\ \dots,\ m$, the set E being closed with respect to disjunction. As it is seen, the set consists of 2^{n+m} images.

Let us consider a situation when an image $e\in E$ is not immediately observable but only after its random distortion of the following simplest type. An image $e:T\rightarrow\{0,1\}$ is transformed into an image $x:T\rightarrow\{0,1\}$ so that at each pixel $t\in T$ a value $e(t)$ is replaced with the value $x(t)=1-e(t)$ with probability α and does not change, $x(t)=e(t)$, with probability $1-\alpha$. For such model it is reasonable to formulate a recognition problem as a maximal likelihood estimation of an ideal image $e\in E$ given result x of its random distortion. At $\alpha<0{,}5$ this problem means to found an ideal image

$$e^*=\arg\min_{e\in E}\sum_{t\in T}\left|e(t)-x(t)\right|. \tag{37}$$

Speaking not formally, here it is necessary to detect a set of vertical lines with a set of horizontal lines in a background, all that being randomly distorted. So, the problem under present consideration can be treated as an extreme simplification of real problems where much more complex objects than strictly vertical lines have to be found, moreover, with much more intricate background than strictly horizontal lines. Yes, the problem under consideration is much more simple than real problems. However, it does not mean that the problem is simple itself, because the set of all possible problems of the type (37) forms an *NP*-complete class.

The problem (37) can be rather naturally formulated as a $(\max,+)$-labelling problem. A set T is a set of pixels, a neighbourhood $\Im$ is a union of a horizontal $\Im^H$ neighbourhood and vertical $\Im^V$ one. A set K consists of four labels: VH, $\bar{V}H$, $V\bar{H}$ and $\bar{V}\bar{H}$ with the following meaning. If $k(t)=VH$ it means that both vertical and horizontal lines pass through the pixel t, if $k(t)=\bar{V}H$ it means that there passes a horizontal line through the pixel, w hile vertical one does not pass. Similar meaning have the labels $V\bar{H}$ and $\bar{V}\bar{H}$.

For pairs $tt'\in\Im^H$ of horizontal neighbours the weights $g_{tt'}(k,k')$ have to be

$$g_{tt'}(VH,VH)=g_{tt'}(\bar{V}H,VH)=g_{tt'}(VH,\bar{V}H)=g_{tt'}(\bar{V}H,\bar{V}H)=$$
$$=g_{tt'}(V\bar{H},V\bar{H})=g_{tt'}(\bar{V}\bar{H},\bar{V}\bar{H})=g_{tt'}(V\bar{H},\bar{V}\bar{H})=g_{tt'}(\bar{V}\bar{H},\bar{V}\bar{H})=0,$$

and $g_{tt'}(k,k')=-\infty$ for each other label pair. Such numbers $g_{tt'}(k,k')$ result in that if it was decided that a horizontal line passes (or does not pass) through some pixel the same decision has to be made about all pixels in the same row. In the same way the weights $g_{tt'}(k,k')$ have to be defined for pairs of vertical neighbours.

Weights $q_t(k)$ have to be

$$\begin{aligned}
&q_t(\bar{V}\bar{H})=0 && \text{if} \quad x(t)=0\\
&q_t(k)=-1 && \text{if} \quad k\neq\bar{V}\bar{H},\ x(t)=0\\
&q_t(\bar{V}\bar{H})=-1 && \text{if} \quad x(t)=1\\
&q_t(k)=0 && \text{if} \quad k\neq\bar{V}\bar{H},\ x(t)=1
\end{aligned}$$

Similarly, like in the previous series of experiments, the aim of present experiments was testing of the algorithms performance. However, $(\max,+)$-problems in previous series have a trivial equivalent for sure, whereas the problems under present consideration are not obligatory equivalent to trivial one. So, there was an additional aim to test how frequently in a random problems population such problems occur, which have no trivial equivalent. The experiments showed that such negative situation never occurred.

The experiments with $(\max,+)$-problems of just described type were fulfilled for images with sizes 10x10, 50x50, 100x100 and 500x500 pixels with probability $\alpha=0{,}1$ of distortion of a pixel signal. Experiments with images of 10x10 pixels were fulfilled with both Algorithms 1 and 2.

It took about 40 sec. on average to solve corresponding problems with Algorithm 1, average number of iteration being about 140. As for the Algorithm 2 the required time was too small to measure it. An average number of iteration was about 10.

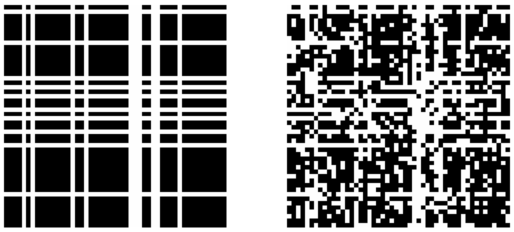
For images with 50x50, 100x100 and 500x500 pixels only Algorithm 2 was used. Experimental results are presented below with three examples, one for each size of image.

Ideal images of some sets of horizontal and vertical lines are shown on fig. 4a, 5a and 6a. Results of their distortion are shown on fig. 4b, 5b and 6b. The distorted images were input to recognition program that formed a corresponding $(\max,+)$-problem, which was solved with the Algorithm 2. The recognition results are not shown while they are exactly the same as ideal images.

It took 1 sec. to process an image of 50x50 pixels (Fig.4b), a number of iterations was 17.

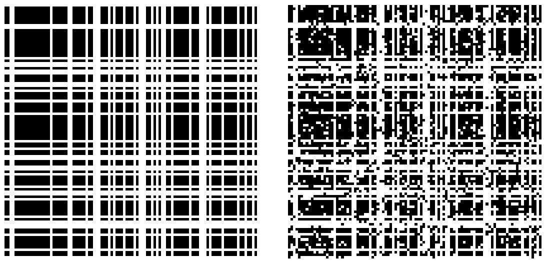
It took 3 sec. to process an image of 100x100 pixels (Fig.5b), a number of iterations was 16.

It took 124 sec. to process an image of 500x500 pixels (Fig.6b), a number of iterations was 36.



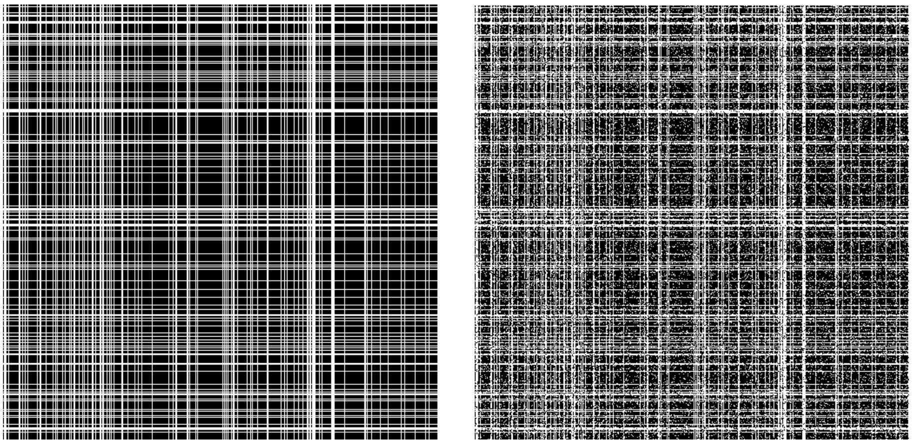
a b

Fig. 4



a b

Fig. 5



a b

Fig. 6

8.5. Image segmentation

Two examples of a wide known texture segmentation problem are presented. An aim and content of this image processing follows from the next generative model [26]. Let T be a field of vision with a conventional for this situation 4-neighbourhood $\Im$. Let $\overline{x}:T\to\{0,1\}$ be a binary image defined on this field and $X^T=\{0,1\}^T$ be a set of all possible images with a set T as their domain. Let on the set X^T some n Gibbs' probability distributions $p^i:X^T\to R$, $i=1,\ 2,\ \dots,\ n$, be defined so that

$$\log p^i\left(\overline{x}\right)=\sum_{tt'\in\Im} g^i_{tt'}\left(x(t),x(t')\right).$$

Let $\overline{k}:T\to\{1,2,\dots,n\}$ be a function referred to as a labelling of T such that $k(t)=i$ means that a pixel t belongs to the segment i. On the set $\{1,2,\dots,n\}^T$ of all possible segmentations a Gibbs' probability distribution $p^\circ:\{1,2,\dots,n\}^T\to R$ is defined so that

$$\log p^\circ\left(\overline{k}\right)=\sum_{tt'\in\Im} g^\circ_{tt'}\left(k(t),k(t')\right).$$

The numbers $g^\circ_{tt'}\left(k,k'\right)$ are chosen here in accordance with a commonly used diagonal model [24]. Namely, for each $tt'\in\Im$

$$g^\circ_{tt'}\left(k,k'\right)=1 \text{ if } k=k',$$

$$g^{\circ}_{tt'}\left(k,k'\right)=0 \text{ if } k\neq k'$$

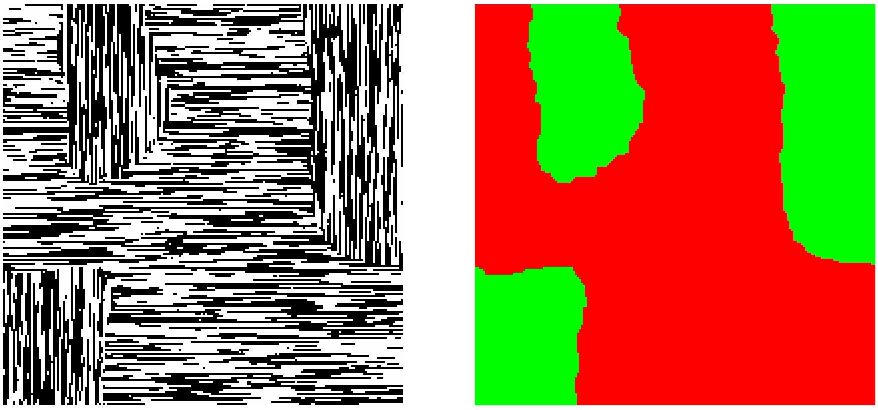
that expresses an intuitive assumption that same labels of two neighbours is more probable than different ones.

Let $\overline{x}_1,\overline{x}_2,...,\overline{x}_n$ be n instances of random images, an image $\overline{x}_i$ being obtained in accordance with its probability distribution p^i. Let $\overline{k}$ be an instance of random segmentation obtained in accordance with p°. An image $\overline{x}:T\rightarrow\left\{0,1\right\}$ is composed of the images $\overline{x}_i$, $i=1,2,...,n$, and the segmentation $\overline{k}$ so that $x\left(t\right)=x_{k(t)}\left(t\right)$. A segmentation problem arises when a generative process is hidden and only the result $\overline{x}$ of this process is observed. The problem consists in that for a given image $\overline{x}$ a reasonable decision has to be made about segmentation $\overline{k}$, which took part in its generation. More strictly, the problem is formulated as a looking for a segmentation $\overline{k}^*$ with maximal quality

$$G\left(\overline{k}\right)=\sum_{tt'\in\Im}g^{\circ}_{tt'}\left(k\left(t\right),k\left(t'\right)\right)+\sum_{t\in T}q^{\circ}_t\left(k\left(t\right),\overline{x}\right). \tag{38}$$

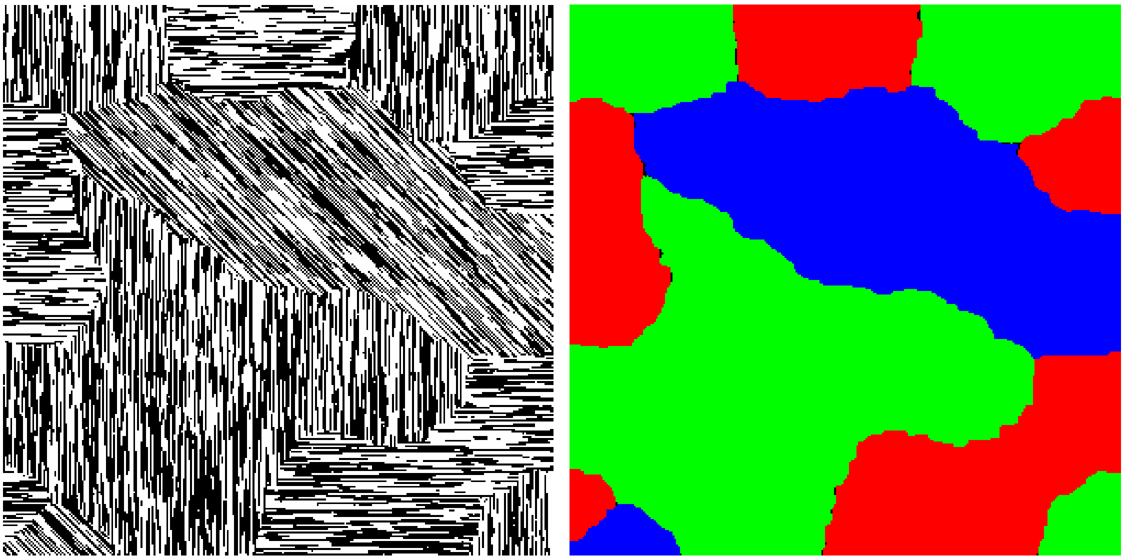
The weights $q^{\circ}_t\left(k\left(t\right),\overline{x}\right)$ were calculated here in a way, which is thoroughly considered in [25] and does not matter in a context of the present paper. It is essential only that if $\left|K\right|=2$ the problem (38) is supermodular independently on weights $q_t\left(k,\overline{x}\right)$ and, consequently, has a trivial equivalent. However, if a number of labels is greater than 2 the problem (38) becomes much more complicated. It is known ([16]) that even for fixed $\left|K\right|=3$ and fixed diagonal weights $g^{\circ}_{tt'}\left(k,k'\right)$ the set of problems of the type (38) with various weights $q_t\left(k\right)$ forms an *NP*-complete class. So, some segmentation problems with $\left|K\right|=3$ have no trivial equivalent. However, such situation has never occurred at experiments.

Two examples of segmentation are shown on Fig. 7 and 8 for $\left|K\right|=2$ and $\left|K\right|=3$. Input images are shown on Fig. 7a and 8a. Their segmentations are shown on Fig. 7b and 8b. The segmentations were obtained as a solution to problem (38) with an Algorithm 2. A size of the image on the Fig. 7a is 200x200, it took 2 sec. and 20 iterations to process it. The image on the Fig. 8a is of size 300x300, a time of its processing was 25 sec., 112 iterations.



a b

Fig. 7. Input image (a) and its segmentation (b)



a b

Fig. 8. Input image (a) and its segmentation (b)

References

1. *M. I. Shlezinger***.** Syntactic analysis of two-dimensional visual signals in the presence of noise // Cybernetics and Systems Analysis, Springer New York, Volume 12, Number 4 / July, 1976. – P. 612-628.

2. *M. I. Schlesinger and B. Flach.* Some solvable subclasses of structural recognition problems. // Czech Pattern Recognition Workshop. – Praha. – 2000. – P. 55-62.

3. *B. Flach.* Strukturelle Bilderkennung. // Technical report, Fakultät Informatik, Technische Universität Dresden, 2002. Habilitation thesis.

4. D. Schlesinger. Strukturelle Ansätze für die Stereorekonstruktion. // Technische Universität at Dresden, Fakultät Informatik, Institut für Künstliche Intelligenz, PhD thesis, July, 2005.

5. *I. Kovtun.* Partial optimal labeling search for a NP-hard subclass of (max,+) problems. // Conf. German Assoc. for Pattern Recognition (DAGM), 2003. - P. 402-409.

6. *V. Kolmogorov, R. Zabih*. What energy functions can be minimized via graph cuts? // 7th Eur. Conf. on Comp. Vision (ECCV), Springer-Verlag, 2002. – P. 65–81.

7. *H. Ishikawa, D. Geiger*. Segmentation by grouping junctions. // IEEE Conf. Comp. Vision and Pattern Recogn. – 1998. – P. 125–131.

8. *H. Ishikawa*. Exact optimization for Markov random fields with convex priors. // IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25. – 2003. - № 10. - P. 1333-1336..

9. *Y. Boykov, O. Veksler, R. Zabih*. Fast approximate energy minimization via graph cuts. // IEEE Trans. Pattern Analysis and Machine Intelligence. – 2001.- **23**.- № 11. - P. 1222-1239.

10. M.*I.Schlesinger, V. Hlavac*. Ten Lectures on Statistical and Structural Pattern Recognition. Computational Imaging and Vision // Kluwer Academic Publishers - Dordrecht / Boston / London., 2002, 520 p.

11. *E. Boros, P.I.Hammer*. Pseudo-Boolean optimization // Discrete Applied Mathematics. - **123**, № 1-3, p.155-225, 2002.

12. *R.E. Burkard, B. Klinz, R. Rudolf.* Perspectives of Monge properties in optimization. // Discrete Applied Mathematics. – 1996. – **70**. - № 2, P.95-161,.

13. *L. Lovasz.* Submodular functions and convexity. // Mathematical Programming – The State of the Art, A. Bachem, M. Grotschel, and B.Korte. EdS. Springer-Verlag, New York, 1983, p. 235-257.

14. *P.L. Hammer.* Some network flow problems solved with pseudo-Boolean programming. //Operations Research. – 1965. – **13**. - P. 388-399.

15. В.К. Коваль, М.И.Шлезингер. Двумерное программирование в задачах анализа изображений. // Автоматика и телемеханика. - М. - 1976 . - № 8. - С. 149-168. *****

16. *M. Wainwright, T. Jaakkola, and A. Willsky.* MAP estimation via agreement on (hyper) trees: message passing and linear programming approaches. // Allerton Conf. on Communication, Control and Computing, 2002.

17. *V. Kolmogorov.* Convergent tree-reweighted message passing for energy minimization. //Microsoft Research. Tech. Rep. MSR-TR-1005-38, 2005.

18. *V. Kolmogorov.* Convergent tree-reweighted message passing for energy minimization. // IEEE Trans. Pattern Analysis and Machine Intelligence. – 2006. – **28**. - № 10. - P. 1568-1583.

19. *T.Werner.* A Linear Programming Approach to Max-sum Problem: // A review. ftp://cmp.felk.cvut.cz/pub/cmp/articles/werner/Werner-TR-2005-25.pdf.

20. М.И. Шлезингер. Математические средства обработки изображений. – Киев, Наукова думка, 1989 – 200 с.

21. Н.З. Шор. Методы минимизации недифференцируемых функций и их приложения. – Киев, Наукова думка, 1979 – 200 с.

22. С.С. Зуховицкий, Л.И. Авдеева. Линейное и выпуклое программирование. М.: Наука. 1967 – 460 с.

23. *Rosenfeld F., Hummel R., Zucker S.* Scene Labeling by Relaxation Operations. // IEEE Transactions on Systems Man and Cybernetics. – 1976.- **SMC-6**.- № 6. – P. 420-433.

24. Potts R. Some generalized order-disorder transformation // Proceedings of the Cambridge Philosophical Society. – 1952. – **48**. -- P. 106–109.

25. Ковтун И.В. Технология текстурной сегментации изображений на основании Марковских случайных полей и решения (мах,+)-задач // УСиМ.—2004. – № 2. – С. 61-66.

26. Ковтун И.В. Текстурная сегментация изображений на основании марковских случайных полей// УСиМ.—2003. – №. 4. – С. 46-55.