Multiresolution Graph Cut Methods in Image Processing and Gibbs Estimation

B. A. Zalesky

1. Plan of Talk

1. Introduction
2. Multiresolution Network Flow Minimum Cut Algorithm
3. Integer Minimization of Modular Functions
4. Integer Minimization of Quadratic Polynomials
5. Combinatorial Minimization of Submodular Functions
   (a) Representation of Functions by Boolean Polynomials
   (b) Modified SFM Algorithm
   (c) Representability of Submogular Boolean Polynomial by Graphs

1. INTRODUCTION

Often, computations to enhance quality of digital pictures, to segment them, to recognize image objects or to find statistical estimates require enormous calculations. In these cases, the real use of the statistical and image estimators is restricted by opportunity to calculate or to evaluate satisfactory their values. Special methods and techniques have been developed to provide that.

Results we present now are based on integer programming and combinatorial optimization methods that were specially proposed to solve practical problems of image processing and statistical Bayes estimation.

In 1996 we studied statistical properties of the Gibbs estimates of images. Simulated annealing (Geman&Geman(1984)) and the heuristic minimum graph cut algorithms by Greig, Porteous and Seheult (1989) were used to find those estimates.

In the paper by Greig, Porteous and Seheult the following problem was formulated: to develop an algorithm that allows computing the Ising estimate of gray scale images by fast optimization methods. It became starting point of our research.

I have pleasure to mention
Olga Veksler,
Yuri Boykov,
Jerome Darbon,
Andrew Goldberg,
Hiroshi Ishikava,
Vladimir Kolmogorov,
Christoph Schnorr,
Marc Sigelle,
whose papers helped me to get my results.

Notations

$S = \{1, \ldots, N\}$	image pixels or graph nodes
$A = \{(i, j),\ i, j \in S\}$	image neighborhood system or graph arcs
$\mathcal{G} = (S, A)$	image graph
$\mathcal{F} = \{f_1, \ldots, f_L\}$	image values or label set of nodes $(f_1 \leq \ldots \leq f_L)$
$\mathbf{I} = \{I_j\}_{j \in S},\ (I_j \in \mathcal{F})$	image or node labels
$\mathcal{M} = \{m_1, \ldots, m_K\}$	estimate values or, in general, another set of labels of nodes $(m_1 \leq \ldots \leq m_K)$

Problems of MAP estimation of images and the statistical Bayes estimation are reduced to the following task: for an image $\mathbf{I}$, which takes a value $\mathbf{f} \in \mathcal{F}$, with the graph $\mathcal{G}$ to determine its estimate

$$\widehat{\mathbf{m}} = \mathrm{argmin}_{\mathbf{m} \in \mathcal{M}^N} U(\mathbf{f}, \mathbf{m}). \tag{1}$$

We propose combinatorial and integer programming techniques that allow computing the estimate (1) for Modular Functions, Quadratic Polynomials and Submodular Functions in a concurrent mode. The algorithms are especially efficient in image processing and the MRF estimation since usually functions $U(\mathbf{f}, \mathbf{m})$ in image processing and the MRF estimation do not have phase transitions (otherwise, changes of neighborhood of object image would completely recolored the object itself).

2. MULTIRESOLUTION NETWORK FLOW MINIMUM CUT ALGORITHM

Let the network $\mathcal{G} = (\widetilde{S}, A)$ consist of $n+2$ numbered *nodes* $\tilde{S} = \{0, 1, 2 \ldots, n, n+1\}$, where $s = 0$ is the *source*, $t = n+1$ is the *sink* and $S = \{1, 2 \ldots, n\}$ are usual nodes. The set of directed *arcs* is $A = \{(i, j) \ : \ i, j \in \tilde{S}\}$. Capacities of arcs are denoted by $d_{i,j} > 0$. Suppose the network $\mathcal{G}$ satisfies the condition:

(G) For every usual node $i \in S$ there is either an arc $(s, i) \in A$ connecting the source s with i or an arc $(i, t) \in A$ connecting i with the sink t but not both of these arcs.

The condition (G) does not restrict the use of the MGMC algorithm. It has been done to simplify writing down notations and proofs. Any network $\mathcal{G}$ can be easily modified to satisfy (G) for $O(n)$ operations (one look through usual network nodes S). The modified network will have the same minimum cut. A network, which corresponds to a binary image, has all white pixels connected with the source and all black ones attached to the sink (see left picture on Fig.1)

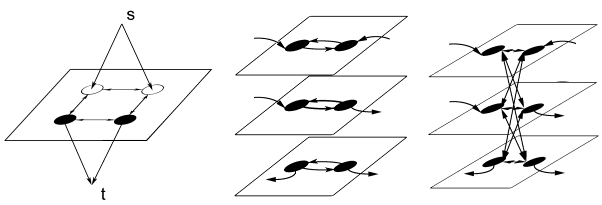

FIGURE 1. Left network satisfies condition (G); Central network corresponds to function $U_{mod,\mathbf{f}}$; Right network presents function $U_{sec,\mathbf{f}}$

Let W, B partition S ($W, B \subset S$, $W \bigcap B = \emptyset$ and $W \bigcup B = S$).

The Boolean vector $\mathbf{x}\ : S \to \{0,1\}^{|S|}$ with coordinates $x_i = 1$, if $i \in W$, and $x_i = 0$ otherwise is the indicator vector of the set W. The set of all such indicator vectors $\mathbf{x}$ is denoted by $\mathbf{X}$.

The *capacity of the* $s - t$ *cut* $C(\mathbf{x})$ is defined as the sum of capacities of all forward arcs going from the set $W \cup \{s\}$ to the set $B \cup \{t\}$, i.e.

$$C(\mathbf{x}) = \sum_{i \in W \cup \{s\}} \sum_{j \in B \cup \{t\}} d_{i,j}.$$

The vector $\mathbf{y}$ with coordinates $y_i = 1$, if $(s,i) \in A$, or $y_i = 0$, if $(i,t) \in A$ (remind that the condition (G) is valid). Set $\lambda_i = d_{s,i}$, if $(s,i) \in A$, or $\lambda_i = d_{i,t}$, if $(i,t) \in A$. The capacities of usual arcs (i,j), $i, j \in S$ will be denoted by $\beta_{i,j} = d_{i,j}$.

The function

$$(2)\quad U(\mathbf{x}) = \sum_{i\in S} \lambda_i(1-2y_i)x_i + \sum_{(i,j)\in S\times S} \beta_{i,j}(x_i - x_j)x_i$$

is connected with $s-t$-cut by $C(\mathbf{x}) = U(\mathbf{x}) + \sum_{i=1}^{n}\lambda_i y_i$. Functions $C(\mathbf{x})$ and $U(\mathbf{x})$ are distinguished only by the constant $\sum_{i=1}^{n}\lambda_i y_i$. Therefore, the solutions $\mathbf{x}^* = \mathrm{argmin}_{\mathbf{x}\in X} U(\mathbf{x})$ entirely identify minimum network flow cuts.

Remark. There is a clear correspondence between networks satisfying condition (G) and binary images. The vector $\mathbf{y}$ can be understood as an original binary image and the vector $\mathbf{x}^*$ as its estimate that minimizes the quadratic function $U(\mathbf{x})$.

For an arbitrary subset of usual nodes $E \subset S$ define two functions

$$(3)\quad U_E(\mathbf{x}) = \sum_{i\in E} \lambda_i(1-2y_i)x_i + \sum_{(i,j)\in(E\times E)\cup(E\times E^c)\cup(E^c\times E)} \beta_{i,j}(x_i - x_j)x_i$$

and

$$V_E(\mathbf{x}) = \sum_{i\in E^c} \lambda_i(1-2y_i)x_i + \sum_{(i,j)\in(E^c\times E^c)} \beta_{i,j}(x_i - x_j)x_i,$$

satisfying the equality $U(\mathbf{x}) = U_E(\mathbf{x}) + V_E(\mathbf{x})$. Also, define restriction $\mathbf{x}_E$ of $\mathbf{x}$ onto the set E, which is the vector $\mathbf{x}_E = (x_i)_{i \in E}$, such that $\mathbf{x} = (\mathbf{x}_E, \mathbf{x}_{E^c})$. The function $V_E(\mathbf{x})$ depends only on $\mathbf{x}_E$

THEOREM 1. *Following properties are valid.*

- **(i):** *If fixed frontier vectors satisfy the condition* $\overset{o}{\mathbf{x}}_{E^c} \leq \overset{o}{\mathbf{z}}_{E^c}$, *then for any solution* $\mathbf{x}'_E = argmin_{\mathbf{x}_E} U(\mathbf{x}_E, \overset{o}{\mathbf{x}}_{E^c})$ *there exists a solution* $\mathbf{z}'_E = argmin_{\mathbf{x}_E} U(\mathbf{x}_E, \overset{o}{\mathbf{z}}_{E^c})$ *such that* $\mathbf{x}'_E \leq \mathbf{z}'_E$.
- **(ii):** *If fixed frontier functions satisfy the condition* $\overset{o}{\mathbf{x}}_{E^c} \leq \overset{o}{\mathbf{z}}_{E^c}$, *then for any solution* $\mathbf{z}'_E = argmin_{\mathbf{x}_E} U(\mathbf{x}_E, \overset{o}{\mathbf{z}}_{E^c})$ *there exists a solution* $\mathbf{x}'_E = argmin_{\mathbf{x}_E} U(\mathbf{x}_E, \overset{o}{\mathbf{x}}_{E^c})$ *such that* $\mathbf{x}'_E \leq \mathbf{z}'_E$.
- **(iii):** *For any frontier condition* $\overset{o}{\mathbf{x}}_{E^c}$ *the set* $\{\mathbf{x}'_E\}_{\overset{o}{\mathbf{x}}_{E^c}}$ *has the minimal (the maximal) element* $\underline{\mathbf{x}}'_E$ $(\overline{\mathbf{x}}'_E)$.
- **(iv):** *If* $\overset{o}{\mathbf{x}}_{E^c} \leq \overset{o}{\mathbf{z}}_{E^c}$, *then* $\underline{\mathbf{x}}'_E \leq \underline{\mathbf{z}}'_E$ *and* $\overline{\mathbf{x}}'_E \leq \overline{\mathbf{z}}'_E$.

Theorem 1 allows to use multiresolution strategy of computing minimum graph cuts. It guarantees existence of two solutions that can be found by known maximum network flow algorithms

$$\mathbf{x}_{0,E} = \mathrm{argmin}_{\mathbf{x}_E} U(\mathbf{x}_E, \mathbf{0}_{E^c})$$

and

$$\mathbf{x}_{1,E} = \mathrm{argmin}_{\mathbf{x}_E} U(\mathbf{x}_E, \mathbf{1}_{E^c}),$$

which satisfy the inequality

$$\mathbf{x}_{0,E} \leq \mathbf{x}'_E \leq \mathbf{x}_{1,E}. \tag{4}$$

Denote the sets on nodes $B = \{k \in E \mid x_{1,E,k} = 0\}$ and $W = \{k \in E \mid x_{0,E,k} = 1\}$. The equalities $\mathbf{0}_B = \mathbf{x}_{1,B} = \mathbf{x}^*_B$ and $\mathbf{1}_W = \mathbf{x}_{0,W} = \mathbf{x}^*_W$ is inferred from (4). If sets B and W are not empty we found coordinates $\mathbf{x}^*_{B \cap W}$ of some solution $\mathbf{x}^*$ of the minimum cut of the original graph.

The main idea of the Multiresolution Minimum Graph Cut algorithm is to partition the set of nodes $S = \bigcup E_i,\ E_i \bigcap E_j = \emptyset$ and then try to find parts of solutions $\mathbf{x}^*_{E_i}$ of $\mathbf{x}^* = \mathrm{argmin}_{\mathbf{x}} U(\mathbf{x})$ operating just with subgraphs that correspond functions $U_{E_i}(\mathbf{x}_{E_i})$.

An appropriate choice of E_i can speed up computations because it allows to reduce number of required operations and also to exploit only the fast cash computer memory. One more advantage of the MGMC algorithm is possibility of its implementation in a concurrent mode.

The algorithm was published in 2000.

FIGURE 2. Original and noisy images and its estimate after execution of 1-st level of the MGMC algorithm

3. INTEGER MINIMIZATION OF MODULAR FUNCTIONS

It was mentioned that Greig, Porteous and Seheult in 1989 posed the problem of exact determining the MAP estimate for the gray-scale Ising model. We present a solution a little bit more general problem, when an image $\mathbf{I}$ is supposes to take values in $\mathcal{F}^N$ and the MAP estimate belongs, in general, to another set of labels $\mathcal{M}^N$.

Denote by

$$(5)\quad U_{mod,\mathbf{f}}(\mathbf{m}) = \sum_{i\in V} \lambda_i |f_i - m_i| + \sum_{(i,j)\in E} \beta_{i,j}|m_i - m_j|, \quad \lambda_i \geq 0, \beta_{i,j} \geq 0.$$

potential function of the generalized gray-scale Ising model and by

$$\widehat{\mathbf{m}}_{mod} = \mathrm{argmin}_{\mathbf{m}\in\mathcal{M}^V} = U_{mod,\mathbf{f}}(\mathbf{m})$$

Let μ, ν be integers and the indicator function $\mathbf{1}_{\left(\mu\geq\nu\right)} = 1$, if $\mu \geq \nu$, and be zero otherwise. For any $\mu \in \mathcal{M}$ and ordered Boolean variables $x(l) =$

$1_{\left(\mu \geq m(l)\right)}$ that satisfy the inequality

$$x(1) \geq x(2) \geq \ldots \geq x(k)$$

the relationship

$$\mu = m(0) + \sum_{l=1}^{K} (m(l) - m(l-1)) x(l), \tag{6}$$

holds true. Vice versa, any non-increasing sequence of Boolean variables $x(1) \geq x(2) \geq \ldots \geq x(k)$ determines integer $\mu \in \mathcal{M}$ according to (6).

Similarly, image values $f_i \in \mathcal{F}$ can be represented as sums

$$f_i = \sum_{\tau=1}^{L} f_i(\tau)$$

of non-increasing sequences of Boolean variables

$$f_i(1) \geq f_i(2) \geq \ldots \geq f_i(L).$$

Let $\mathbf{x}(l) = (x_1(l), x_2(l), \ldots, x_N(l))$, $l = 1, \ldots, K$ be Boolean vectors. Vectors $\mathbf{z}(l)$ are with coordinates

$$z_i(l) = \frac{1}{m(l) - m(l-1)} \sum_{\tau=m(l-1)+1}^{m(l)} f_i(\tau). \tag{7}$$

$$(i = 1, \ldots, N, \; l = 1, \ldots K)$$

and

$$|\mathbf{x}| = \sum_{i=1}^{N} |x_i|$$

is the norm of the vector. The following Proposition is valid.

PROPOSITION 2. *For all integer $\nu \in \mathcal{M}$ and all $f_i \in \mathcal{F}$*

$$(8)\quad |\nu - f_i| = \left| m(0) - \sum_{\tau=1}^{m(0)} f_i(\tau) \right| + \sum_{l=1}^{K} \left| (m(l) - m(l-1))\mathbf{1}_{\left(\nu \geq m(l)\right)} - \sum_{\tau=m(l-1)+1}^{m(l)} f_i(\tau) \right| + \sum_{\tau=m(k)+1}^{L} f_i(\tau)$$

Thus, for any image value $\mathbf{f} \in \mathcal{F}^N$ and any value $\mathbf{m} \in \mathcal{M}^N$ of the estimate the function $U_{mod,\mathbf{f}}(\mathbf{m})$ can be represented in the following form

$$(9)\quad U_{mod,\mathbf{f}}(\mathbf{m}) = \sum_{l=1}^{K} (m(l) - m(l-1))u(l, \mathbf{x}(l)) + const,$$

where const $= N\left|m(0) - \sum_{\tau=1}^{m(0)} f_i(\tau)\right|$, *$N$-dimensional Boolean vectors*

$$\mathbf{x}(l) = (1_{\left(m_1 \geq m(l)\right)}, 1_{\left(m_2 \geq m(l)\right)}, \ldots, 1_{\left(m_n \geq m(l)\right)})$$

and functions

$$u(l, \mathbf{b}) = \sum_{i \in S} \lambda_i |z_i(l) - b_i| + \sum_{(i,j) \in A} \beta_{i,j} \left|b_i - b_j\right|.$$

for N-dimensional Boolean vectors $\mathbf{b}$.

Denote by

$$\text{(10)} \qquad \check{\mathbf{x}}(l) = \text{argmin}_{\mathbf{b}} u(l, \mathbf{b}), \quad (l = 1, \ldots, K)$$

the Boolean solutions that minimize $u(l, \mathbf{b})$ and note that

$$\mathbf{z}(1) \geq \mathbf{z}(2) \geq \ldots \geq \mathbf{z}(K).$$

Moreover,

$$z_i(1) = z_i(2) = \ldots = z_i(\nu - 1) = 1,$$
$$0 \leq z_i(\nu) \leq 1, z_i(\nu + 1) = \ldots = z_i(K) = 0$$

for some integer $1 \leq \nu \leq K$. It is easy to see that solutions $\check{\mathbf{x}}(l)$ are in general unordered in a sense that for $l_l < l_2$ unordered solutions $\check{\mathbf{x}}_{l_1} \not\geq \check{\mathbf{x}}_{l_2}$ can exist. Nevertheless, there always exists non-increasing

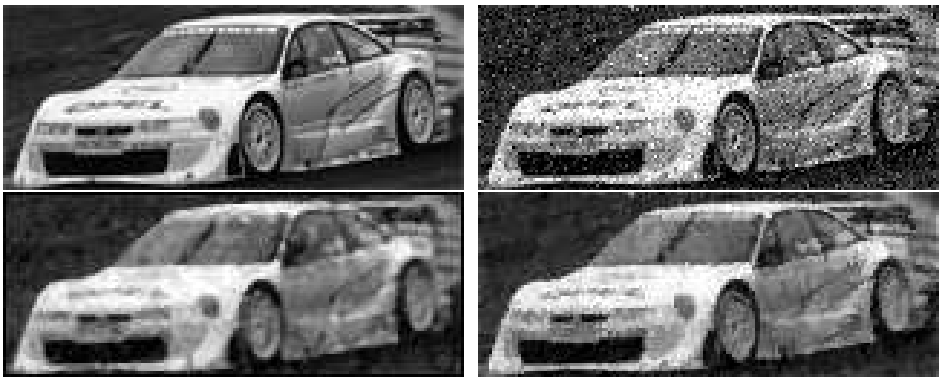

FIGURE 3. Original, noisy, filtered by median and $\widehat{\mathbf{m}}_{mod}$ -estimate images

sequence

$$\check{\mathbf{x}}(1) \geq \check{\mathbf{x}}(2) \geq \ldots \geq \check{\mathbf{x}}(K)$$

of solutions of problem (10).

This decomposition, in fact, allows computation of $\widehat{\mathbf{m}}_{mod}$ as

$$\widehat{\mathbf{m}}_{mod} = \mathbf{m}(0) + \sum_{l=1}^{K}(m(l) - m(l-1))\check{\mathbf{x}}(l)$$

in a concurrent mode with the use of solution $\check{\mathbf{x}}(l-1)$ to compute $\check{\mathbf{x}}(l)$.

The result was published in 2001.

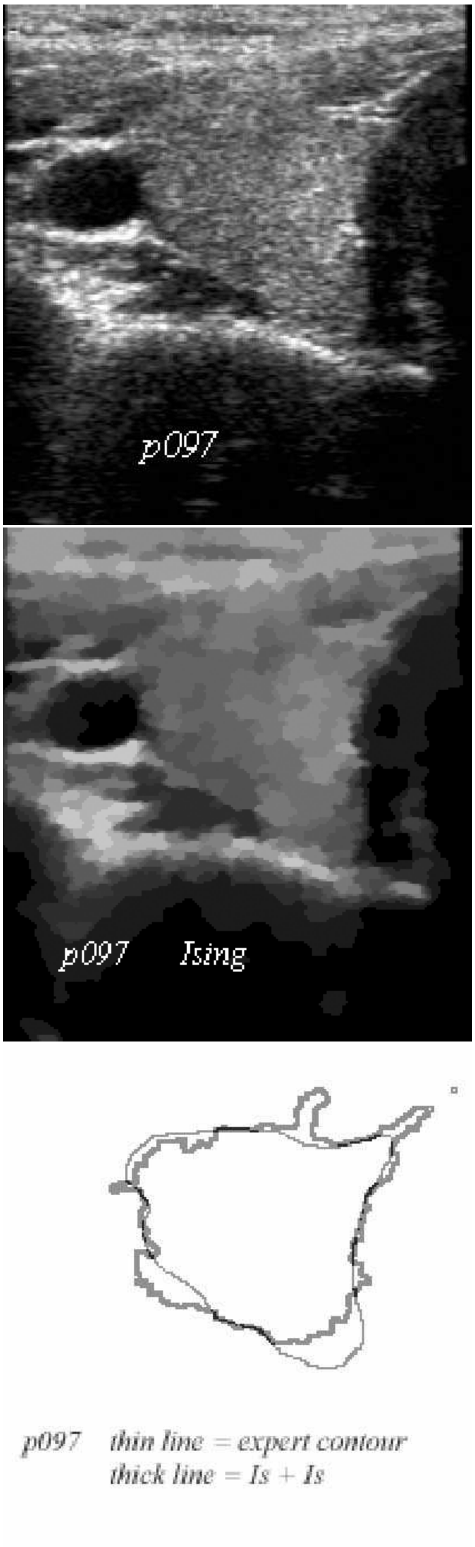

FIGURE 4. Ultrasonic image of thyroid gland, clusterized and then segmented by the gray-scale modular function

4. INTEGER MINIMIZATION OF SQUARE POLYNOMIALS

Submodular square polynomials on integral vectors $\mathbf{f} \in \mathcal{F}^N$, ($\mathcal{F} = \{f_1, \ldots, f_L\}$) and $\mathbf{m} \in \mathcal{M}^N$, ($\mathcal{M} = \{m_1, \ldots, m_K\}$) can be written into the form

$$U_{sec,\mathbf{f}}(\mathbf{m}) = \sum_{i \in V} \lambda_i (f_i - m_i)^2 + \sum_{(i,j) \in E} \beta_{i,j} (m_i - m_j)^2, \quad \lambda_i \geq 0, \beta_{i,j} \geq 0$$

and after represented by the Boolean polynomial

$$U_{sec,\mathbf{f}}(\mathbf{m}) = P(\mathbf{x}(1), \ldots, \mathbf{x}(K)), \mathbf{m} = \sum_{l=1}^{K} \mathbf{x}(l),$$

which is non-submodular.

Submodular Boolean polynomial $Q(\mathbf{x}(1), \ldots, \mathbf{x}(K))$ is constructed so that

$$Q(\mathbf{x}(1), \ldots, \mathbf{x}(K)) \geq P(\mathbf{x}(1), \ldots, \mathbf{x}(K))$$

and their points of minima coinside

$$(\mathbf{q}^*(1), \mathbf{q}^*(2), \ldots, \mathbf{q}^*(K)) = \operatorname{argmin}_{\mathbf{x}(1),\mathbf{x}(2),\ldots,\mathbf{x}(K)} Q(\mathbf{x}(1), \ldots, \mathbf{x}(K)) = \operatorname{argmin}_{\mathbf{x}(1),\mathbf{x}(2),\ldots,\mathbf{x}(K)} P(\mathbf{x}(1), \ldots, \mathbf{x}(K)).$$

Besides, any set of Boolean vectors minimizing Q turns to be decreasing:

$$\mathbf{q}^*(1) \geq \mathbf{q}^*(2) \geq \ldots \geq \mathbf{q}^*(K).$$

Therefore each solution $\widehat{\mathbf{m}}_{sec}$ is given by the formula $\widehat{\mathbf{m}}_{sec} = \sum_{l=1}^{K} \mathbf{q}^*(l)$.

The result was published in 2001.

5. COMBINATORIAL MINIMIZATION OF SUBMODULAR FUNCTIONS

Representation of Functions by Boolean Polynomials. In general, any function $U(\mathbf{x})$ depending on variables $\mathbf{x} = (x_1, \dots, x_N)$, which take values in an arbitrary finite totally ordered set $\mathcal{R}^N = \{r_1, r_2, \dots, r_K\}^N, \ r_1 \le r_2 \le \dots \le r_K$, can be represented by a Boolean polynomial. However, for the sake of simplicity we consider instead of $U(\mathbf{x})$ the function

$$V(\mathbf{j}) = U(r_{j_1}, r_{j_2}, \dots, r_{j_N})$$

on integer variables $\mathbf{j} = (j_1, j_2, \dots, j_N)$ with coordinates $j_l \in \{1, 2, \dots, K\} = \mathcal{N}_K$.

To exploit the SFM we represent once more the integer variables j_i as the sum of ordered Boolean ones

$$j_i = \sum_{l=1}^{K} x_i(l), \tag{11}$$

where $x_i(l) \in \{0, 1\}, \ x_i(1) \ge x_i(2) \ge \dots \ge x_i(K)$ and after decomposse the function $V(\mathbf{j})$ by a Boolean polynomial.

Remind that the Boolean vector $\mathbf{x}(l)$ is of the form $(x_1(l), x_2(l), \dots, x_N(l))$, and for two vectors $\mathbf{x}, \mathbf{z}$ we say $\mathbf{x} \geq \mathbf{z}$, if $x_i \geq z_i$, $i = 1, \ldots, N$ (respectively, $\mathbf{x} \not\geq \mathbf{z}$, if there is a couple of indexes i, j such that $x_i < z_i$ and $x_j > z_j$). Then

$$V(\mathbf{j}) = V\left(\sum_{l=1}^{K} \mathbf{x}(l)\right), \quad \mathbf{x}(1) \geq \mathbf{x}(2) \ldots \geq \mathbf{x}(K).$$

The following statement is valid.

PROPOSITION 3. *The expansion of* $V(\mathbf{j})$ *into a polynomial in ordered Boolean variables* $\mathbf{x}(1) \geq \mathbf{x}(2) \geq \ldots \geq \mathbf{x}(K)$ *is of the form*

$$(12) \quad V(\mathbf{j}) = \widetilde{P}_V\left(\mathbf{x}(1), \mathbf{x}(2) \ldots, \mathbf{x}(K)\right) = V(\mathbf{0}) + \sum_{m=1}^{N} \sum_{1 \leq l_1 < \ldots < l_m \leq N} \sum_{\mu_1, \ldots, \mu_m = 1}^{K} \Delta_{l_1, \ldots, l_m} V\left(\sum_{\tau=1}^{m} \mu_\tau \mathbf{e}_{l_\tau}\right) \times \prod_{\tau=1}^{m} x_{l_\tau}(\mu_\tau).$$

where

$$\Delta_i V(\mathbf{j}) = V(\mathbf{j}) - V(\mathbf{j} - \mathbf{e}_i),$$

$$\mathbf{e}_i = (\overbrace{\underbrace{0, \ldots, 0}_{i-1}, 1, 0, \ldots, 0}^{N}), \quad \mathbf{j} \in \mathcal{N}_K$$

The expansion (12) does not allow to turn immediately to the problem of the Boolean minimization because of ordered variables. To avoid this disadvantage we will use another polynomial

$$P_V\left(\mathbf{x}(1), \mathbf{x}(2) \ldots, \mathbf{x}(k)\right) = \widetilde{P}_V\left(\mathbf{x}(1), \mathbf{x}(2) \ldots, \mathbf{x}(k)\right)$$
$$+ \; const \sum_{\mu=1}^{N} \sum_{l=2}^{K} (x_\mu(l) - x_\mu(l-1)) x_\mu(l),$$

for sufficiently large constant $const > 0$. This polynomial satisfy the following properties.

Proposition 4. *The inequality $\widetilde{P}_V \le P_V$ holds true. Any collection of Boolean vectors $\mathbf{z}^* = (\mathbf{x}^*(1), \mathbf{x}^*(2) \ \ldots, \mathbf{x}^*(k))$ that minimizes P_V, i.e.*

$$\mathbf{z}^* = \arg\min_{\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(k)} P_V\left(\mathbf{x}(1), \mathbf{x}(2) \ldots, \mathbf{x}(k)\right)$$

turns to be ordered (non-increasing). Therefore, this collection minimizes the polynomial $\widetilde{P}_V$, and the integer vector $\mathbf{j}^* = \sum_{l=1}^{K} \mathbf{x}^*(l)$ minimizes the original function $V(\mathbf{j})$.

Hence, instead of minimization of $V(\mathbf{j})$ we can consider the problem of Boolean minimization of submodular polynomials

$$P_V(\mathbf{z}) = a_0 + \sum_{m=1}^{N} \sum_{1 \le l_1 < \ldots < l_m \le N} a_{l_1, l_2, \ldots, l_m} \prod_{i=1}^{m} z_{l_i},$$
$$\dim(\mathbf{z}) = KN.$$

However, we should remember that although the problem of minimization of submodular Boolean polynomials admits solution of polynomial complexity on input data, the polynomial $P_V(\mathbf{z})$, in the general case, consists of exponential number $O(e^N)$ of monomials. Therefore, to guarantee existence of fast minimizing algorithms we have to use for applications $P_V(\mathbf{z})$ having polynomial number $O(N^d)$ of summands. Among them are, for example, polynomials of a fixed degree and polynomials with Markov type dependence of variables.

Modified SFM Algorithm. For two Boolean vectors $\mathbf{x}, \mathbf{y}$ let us denote by $\mathbf{x} \vee \mathbf{y}$ (respectively, by $\mathbf{x} \wedge \mathbf{y}$) the vector with coordinates $\max(x_i, y_i)$ (respectively, with coordinates $\min(x_i, y_i)$).

Definition 1. *A function f is called **submodular** if for any Boolean vectors $\mathbf{x}, \mathbf{y}$ it satisfies the* *inequality*

$$f(\mathbf{x} \wedge \mathbf{y}) + f(\mathbf{x} \vee \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}).$$

Remind that the set $S = \{1, \ldots, N\}$ and that for any $\mathcal{D} \subset S$ the set $\mathcal{D}^c = S \backslash \mathcal{D}$. The vector $\mathbf{x}_{\mathcal{D}}$ with coordinated $x_i, \ i \in \mathcal{D}$ is the restriction of $\mathbf{x}$ to the set $\mathcal{D}$, the abbreviation $P(\mathbf{x}_{\mathcal{D}}) = P(\mathbf{x}_{\mathcal{D}}, \mathbf{0}_{\mathcal{D}^c})$.

Below we always suppose for simplicity that $P(\mathbf{0}) = 0$.

The polynomial $P(\mathbf{x})$ can be represented into the form

$$P(\mathbf{x}) = P\{\mathcal{D}\}(\mathbf{x}) + P(\mathbf{x}_{\mathcal{D}^c}), \quad \mathcal{D} \subset S,$$

where each monomial of $P\{\mathcal{D}\}(\mathbf{x})$ contains at least one variable $x_i,\ i \in \mathcal{D}$ (and, in general, it depends on $\mathbf{x}_{\mathcal{D}^c}$), and $P(\mathbf{x}_{\mathcal{D}^c})$ is independent of $\mathbf{x}_{\mathcal{D}}$. Denote by

$$\mathcal{M}(\mathring{\mathbf{x}}_{\mathcal{D}^c}) = \left\{ \arg\min_{\mathbf{x}_{\mathcal{D}}} P\{\mathcal{D}\}(\mathbf{x}_{\mathcal{D}}, \mathring{\mathbf{x}}_{\mathcal{D}^c}) \right\}$$

the set of minima of the polynomial $P\{\mathcal{D}\}$ on $\mathbf{x}_{\mathcal{D}}$ for fixed $\mathring{\mathbf{x}}_{\mathcal{D}^c}$. Then the following result is valid.

COROLLARY 5. *Let* $\mathring{\mathbf{x}}_{\mathcal{E}^c} \leq \mathring{\mathbf{z}}_{\mathcal{E}^c}$ *be any ordered pair of vectors, then:*

- **(i):** *for any* $\mathbf{x}'_{\mathcal{D}} \in \mathcal{M}(\mathring{\mathbf{x}}_{\mathcal{D}^c})$ *there exists a solution* $\mathbf{z}'_{\mathcal{D}} \in \mathcal{M}(\mathring{\mathbf{z}}_{\mathcal{D}^c})$ *such that* $\mathbf{x}'_{\mathcal{D}} \leq \mathbf{z}'_{\mathcal{D}}$.
- **(ii):** *For any* $\mathbf{z}'_{\mathcal{D}} \in \mathcal{M}(\mathring{\mathbf{z}}_{\mathcal{D}^c})$ *there exists a solution* $\mathbf{x}'_{\mathcal{D}} \in \mathcal{M}(\mathring{\mathbf{x}}_{\mathcal{D}^c})$ *such that* $\mathbf{x}'_{\mathcal{D}} \leq \mathbf{z}'_{\mathcal{D}}$.
- **(iii):** *The sets* $\mathcal{M}(\mathring{\mathbf{x}}_{\mathcal{D}^c})$ *and* $\mathcal{M}(\mathring{\mathbf{z}}_{\mathcal{D}^c})$ *have minimal* $\underline{\mathbf{x}}'_{\mathcal{D}}, \underline{\mathbf{z}}'_{\mathcal{D}}$ *and maximal* $\overline{\mathbf{x}}'_{\mathcal{D}}, \overline{\mathbf{z}}'_{\mathcal{D}}$ *elements.*
- **(iv):** *The minimal and maximal elements are ordered, i.e.* $\underline{\mathbf{x}}'_{\mathcal{D}} \leq \underline{\mathbf{z}}'_{\mathcal{D}}$ *and* $\overline{\mathbf{x}}'_{\mathcal{D}} \leq \overline{\mathbf{z}}'_{\mathcal{D}}$.

Corollary 5 allows the use of multiresolution approach to find minima of submodular Boolean polynomial. The approach is similar to the Multiresolution Graph Minimum Cut algorithm.

Representability of Submogular Boolean Polynomial by Graphs.

In paper by Kolmogorov&Zabih (2002) the following definition of graph representability is given.

Definition 2. *A function $U(\mathbf{x})$ of N Boolean variables is called graph representable if there exists a network $\widetilde{\mathcal{G}} = (\widetilde{\mathcal{S}}, \widetilde{\mathcal{A}})$ with number of nodes $M \geq N$ and the cost function $C(\mathbf{z})$ such that for some constant c*

$$U(\mathbf{x}) = \min_{z_{N+1},\ldots,z_M} C(x_1, \ldots, x_N, z_{N+1}, \ldots, z_M) + c.$$

Obviously, there is another definition of graph representability.

Definition 3. *We say a function $U(\mathbf{x})$ of N Boolean variables can be represented by a graph if there exists a submodular quadratic polynomial $\mathbf{p}(\mathbf{z})$ of M $(M \geq N)$ on Boolean variables, which satisfies the equality*

$$U(\mathbf{x}) = \min_{z_{N+1},\ldots,z_M} \mathbf{p}(x_1, \ldots, x_N, z_{N+1}, \ldots, z_M) + c.$$

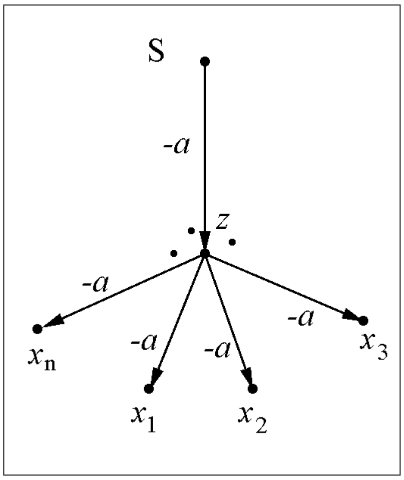

FIGURE 5. Graph representing $ax_1x_2\ldots x_m$, $a<0$

The following Lemma allows to represent Boolean polynomials, which have all coefficients of nonlinear monomials negative, by graphs.

LEMMA 6. *For any natural m and real $a<0$ the Boolean polynomial $ax_1x_2\ldots x_m$ is graph representable.*

The monomials of order greater than 2 with positive coefficients can not be represented by graphs directly. To do it we should accompany those positive monomials by quadratic items with large enough negative coefficients. It requires additional nodes in graphs that represent the monomials.

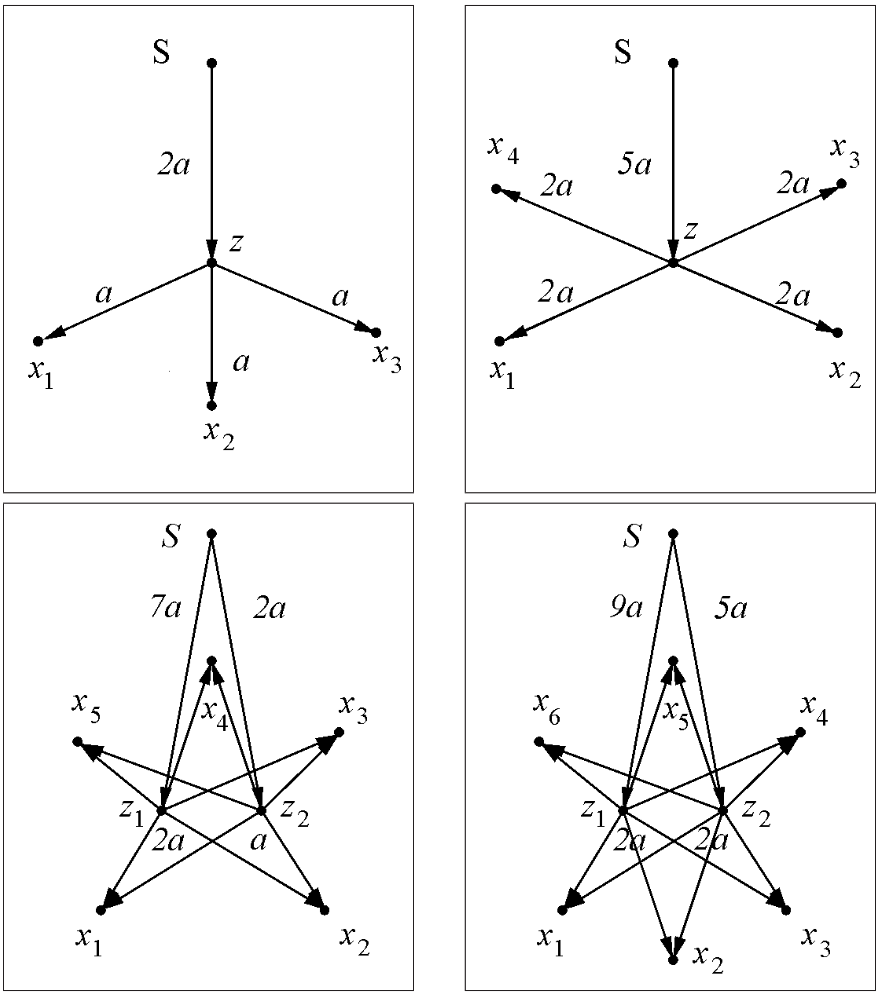

FIGURE 6. Graphs representing $ax_1x_2 \dots x_m - a\sum_{1 \le i < j \le m} x_ix_j$, $a > 0$, $m = 3, 4, 5, 6$

LEMMA 7. *For any natural $m > 2$ and $a > 0$ the Boolean polynomial* $P(x_1, x_2, \dots, x_N) = ax_1x_2 \dots x_m - a\sum_{1 \le i < j \le m} x_ix_j$ *is graph representable.*

Lemmas 6,7 allows to formulate

THEOREM 8. *The following Boolean polynomials P are graph representable:*

(i): *Polynomials with all coefficients of non-linear monomials negative;*

(ii): *Submodular polynomials having all coefficients before monomials of order 3 and high positive;*

(iii): *Polynomials positive coefficients of which $b^{+}_{l_1,l_2,\dots,i,\dots,j,\dots,l_m} > 0$ satisfy the condition*

$$a_{i,j} + \sum_{m=1}^{n-2} \sum_{l_1<l_2<\dots<l_m} b^{+}_{l_1,l_2,\dots,i,\dots,j,\dots,l_m} \le 0, \quad \forall i,j \in S.$$

The results were published in 2003.

References

[1] Zalesky B.A. *Multiresolution Minimum Network Cut Algorithm and its Application to Image Recovering.* Proceedings of the International workshop "Discrete optimization Methods in Scheduling and Computer-aided Design" (Minsk, September 5-7, 2000), p. 173-187.

[2] Zalesky B.A. *Computation of Gibbs estimates of gray-scale images by discrete optimization methods.* Proceedings of the Sixth International Conference PRIP'2001 (Minsk, May 18-20, 2001), p. 81-85.

[3] Zalesky B.A. *Efficient integer-valued minimization of quadratic polynomials with the quadratic monomials* $b_{i,j}^2(x_i - x_j)^2$. Dokl. NAN Belarus, 45(2001), No. 6, p.9-11. (In Russian)

[4] Zalesky B.A. *Gibbs Classifiers with Submodular Energy Functions.* Abstracts of Barcelona Conference on Asymptotic Statistics BAS2003 (Barcelona, September 2-6, 2003). p. 49-50.

[5] Zalesky B.A. *Network Flow Optimization for Restorations of Images.* Journal of Applied Mathematics, 2(2002), No. 4, p. 199-218.

[6] Zalesky B.A. *Gibbs Classifiers.* Probability Theory and Mathematical Statistics 70(2004). p. 36-46.

[7] Zalesky B.A. *Integer Programming Methods in Image Processing and Bayes Estimation.* In: Soft Computing in Image Processing - Recent Advances Series: Studies in Fuzziness and Soft Computing , Vol. 210. Nachtegael, M.; Van der Weken, D.; Kerre, E.E.; Philips, W. (Eds.) 2007, XXI, 500 p., ISBN-10: 3-540-38232-1, ISBN-13: 978-3-540-38232-4.