Application of information-theoretical performance criterion to image segmentation

Murashov D.M.

FRC CSC RAS, Moscow, Russia
d_murashov@mail.ru

This research is funded by the Russian Foundation for Basic Research, grants Nos. 15-07-09324 and 15-07-07516.

11-th International Conference
"Intelligent Data Processing:Theory and Applications 2016“
October 9-14, 2016
Barcelona, Spain

Problem Formulation

Segmentation system: $V = F(U, \delta)$,

where U, V are the input and output images; δ is a parameter,

$$U : R^2 \to R; V : R^2 \to R; \delta \in R, F : R \times R^2 \to R.$$

Input image U generates a set of Q segmented images

$$\mathfrak{V} = \{V_1, V_2, ..., V_Q\}.$$

It is necessary to find $q_{\min} = \arg\min_q \left(M(U, V_q) \right), \; q = 1, 2, ..., Q,$

where $M(U, V_q)$ is a quality measure.

Conventional Techniques

1. Comparing results of segmentation with an image segmented by an expert and accepted as a groundtruth (Arbelaez, 2011).
2. Considering segmentation operation as clustering of pixels:
   - set-theoretical measures;
   - statistical measures;
   - information-theoretical measures.

The most commonly used are:

- chi-square measure;
- Rand Index;
- Fowlkes-Mallows measure;
- mutual information;
- variation of information.

Standard methodology for estimating efficiency of segmentation algorithms is not yet developed.

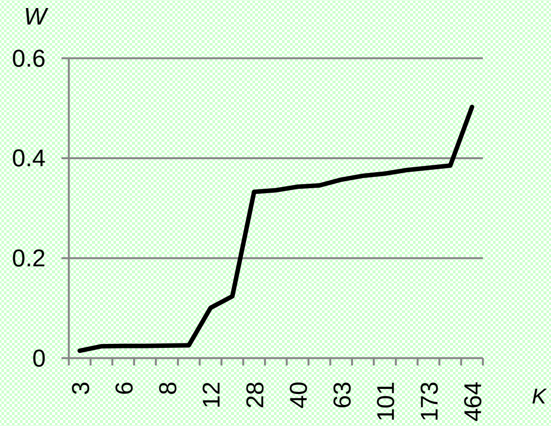
Conventional Techniques

Another approach: estimation of similarity between the segmented and original images.

Similarity measure:

• weighted uncertainty index.

(I. Frosio, E.R. Ratner. Adaptive Segmentation Based on a Learned Quality Metric. Proc. VISAPP 2015).



Atick, J., Norman, A., 1990:

• theoretical-information model of the human visual system.

The main principle: minimizing data redundancy at the early stages of the signal processing in the human visual system.

(Atick, J., Norman, A., 1990. Towards a theory of early visual processing. Neural Computation archive 2(3):308–320.)

Choosing the Best Segmentation

Model of segmentation system: $V = F(U+\eta)$,

where U, V are the input and output images; η is noise;

F is a transformation function; U, V are the continuous random variables.

η is a Gaussian random variable with $M_\eta = 0$ and variance σ_η^2 ;

$\operatorname{cov}(\eta, V) = 0.$

We propose:

criterion of the segmentation quality: $R = 1 - \frac{I(U;V)}{C(V)}$,

where R is the redundancy measure; .

$I(U;V)$ is the mutual information; $C(V)$ is the channel capacity.

We define $C(V) = H(V)$, where $H(V)$ is the entropy of the output.

$$R = 1 - \frac{I(U;V)}{H(V)} = \frac{H(V|U)}{H(V)},$$

where $H(V|U)$ is the conditional entropy.

Choosing the Best Segmentation

Probability mass function of the output: $p(v)=\sum_{k=1}^{K}P(v_k)\delta(v-v_k),$

where $P(v_k)$ is a probability of lightness value, $P(v_k)=\dfrac{1}{K}$;

$\delta(v-v_k)$ is a delta-function; K is number of segments.

Differential entropy of V : $H(V)=-\int_{-\infty}^{+\infty}p(v)\log p(v)dv=\log K.$

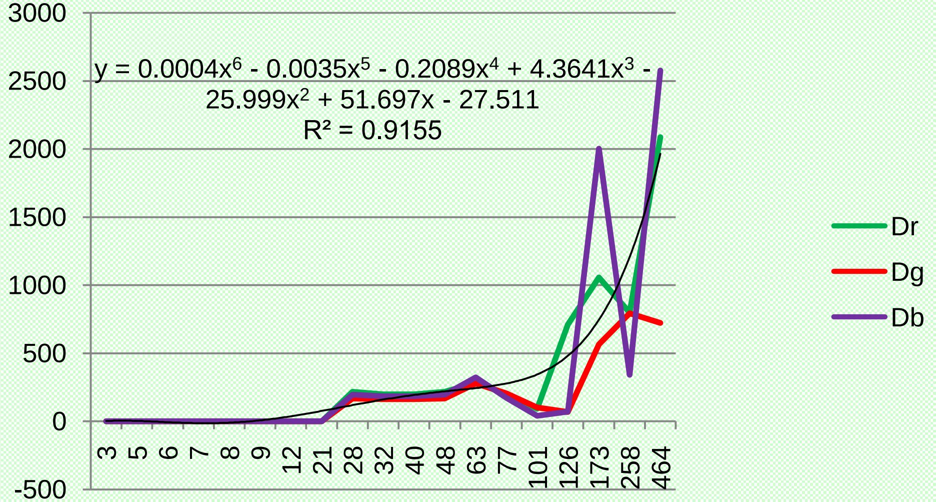
Conditional entropy: $H(V|U)=H(\eta).$

Differential entropy of noise: $H(\eta)=\dfrac{1}{2}\left[\log e+\log(2\pi\sigma_\eta^2)\right].$

Redundancy measure: $R=\dfrac{\log e+\log(2\pi\sigma_\eta^2)}{2\log K}.$

Choosing the Best Segmentation

Redundancy measure: $$R = \frac{\log e + \log(2\pi\sigma_\eta^2)}{2\log K}.$$



SLIC Segmentation Algorithm (R. Achanta et al., 2012)

Point of Image I: $p = (c_1, c_2, c_3, x, y)^T.$

1. Partitioning image into K fragments of size $a \times a$ with centers C_i
2. Moving the centers C_i to seed locations corresponding to lowest gradient position $\nabla I(C_i) \to \min$.
3. Forming clusters in the center neighborhood of size $2a \times 2a$.

   Distance between p and C_i :

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{a}\right)^2 m^2}, \qquad d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2},$$

$$d_c = \sqrt{(c_{1j} - c_{1i})^2 + (c_{2j} - c_{2i})^2 + (c_{3j} - c_{3i})^2}.$$

4. Computing new centers $C_i^{'}$.
5. Repeat steps 3 and 4 are until reaching prescribed precision.

Postprocessing Procedure

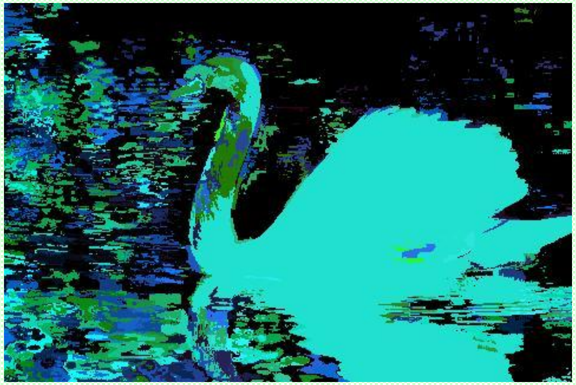
1. Merging neighboring superpixels.

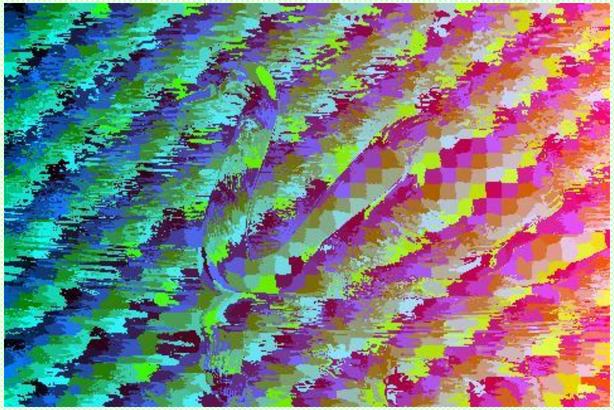
$$\Im_{ij} = \Im_i \bigcup \Im_j \text{, if } d_c(C_i, C_j) < \delta_1,$$

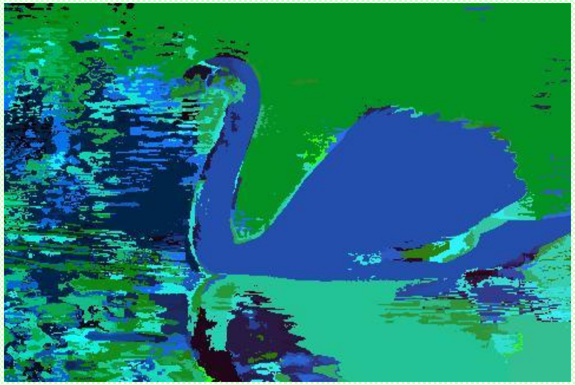
$$d_c(C_i, C_j) = \sqrt{(c_{1j} - c_{1i})^2 + (c_{2j} - c_{2i})^2 + (c_{3j} - c_{3i})^2}.$$

2. Merging superpixels in the whole image.

$$\Im_{ij} = \Im_i \bigcup \Im_j \text{, if } d_c(C_i, C_j) < \delta_2.$$







Computing Experiment

Tasks of the experiment

1. Generating set of segmented images $\mathfrak{V} = \{V_1, V_2, \dots, V_Q\}$ from input U and computing R_W :

$$R_W = \frac{R_L H_L(U) + R_a H_a(U) + R_b H_b(U)}{H_L(U) + H_a(U) + H_b(U)};$$

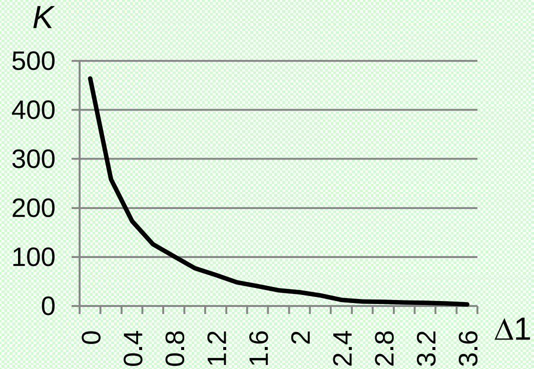
2. Estimating segmentation quality :

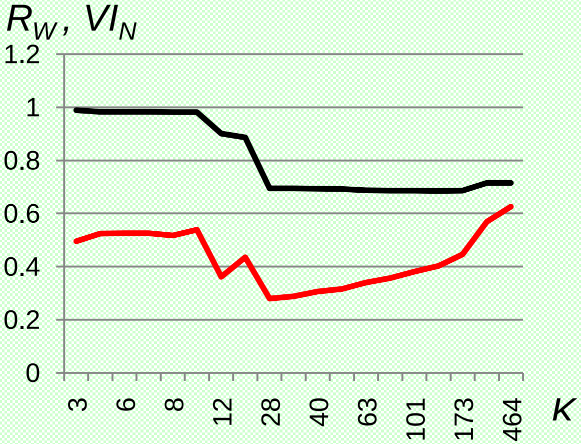
$$VI_W(U, V_q) = \frac{VI_L H_L(U) + VI_a H_a(U) + VI_b H_b(U)}{H_L(U) + H_a(U) + H_b(U)},$$

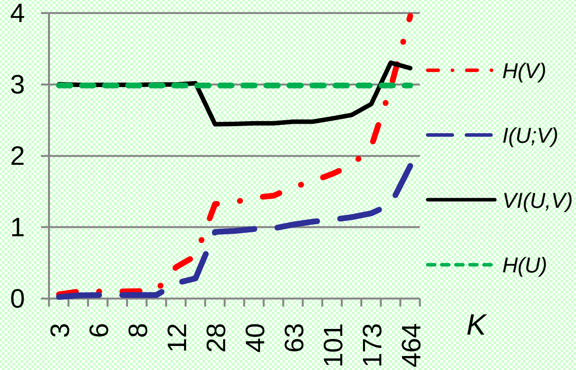
$$VI_i(U, V_q) = H_i(U) + H_i(V_q) - 2I_i(U, V_q); \qquad VI_N = \frac{VI(U,V)}{H(U,V)}$$

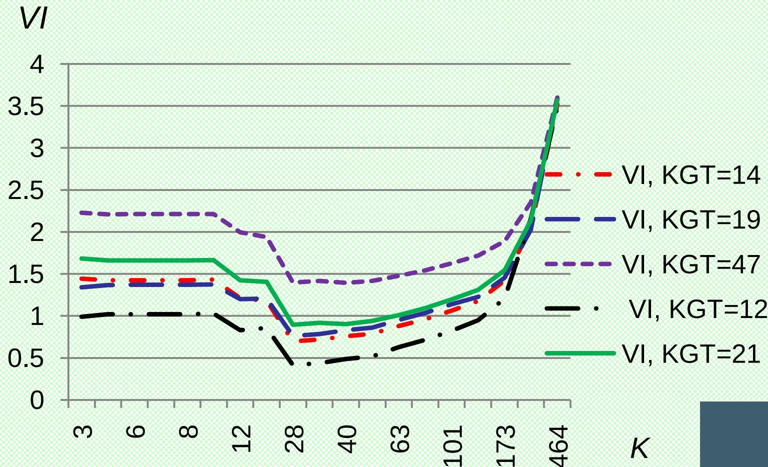
3. Comparing segmented images $V_q,\ q = 1, 2, \dots, Q$ with the groundtruth segmentations $V_t^{GT},\ t = 1, 2, \dots, T.$

Computing Experiment: Results









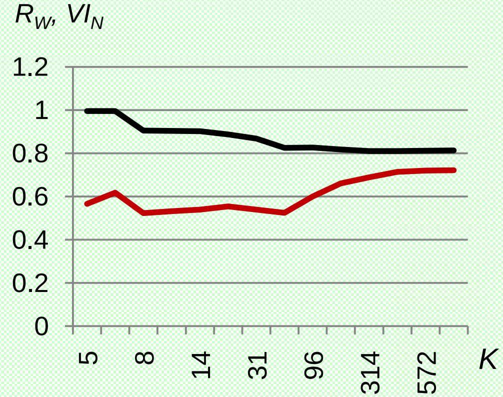
$$VI = H(U) + H(V) - 2I(U;V)$$

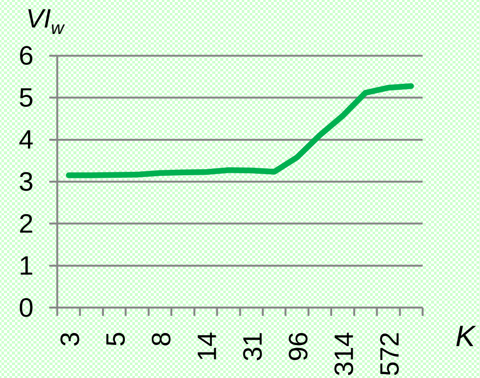
$$VI_N = \frac{VI(U,V)}{H(U,V)}$$

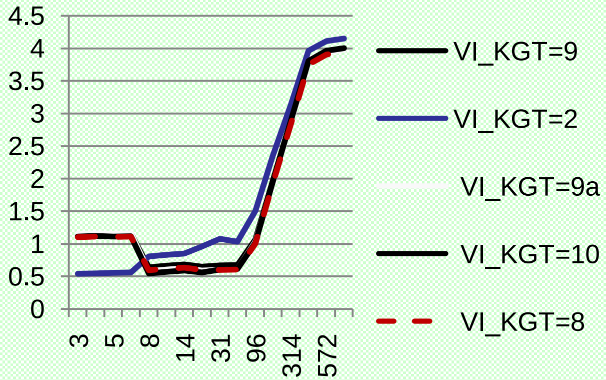
Segmented image, K = 28

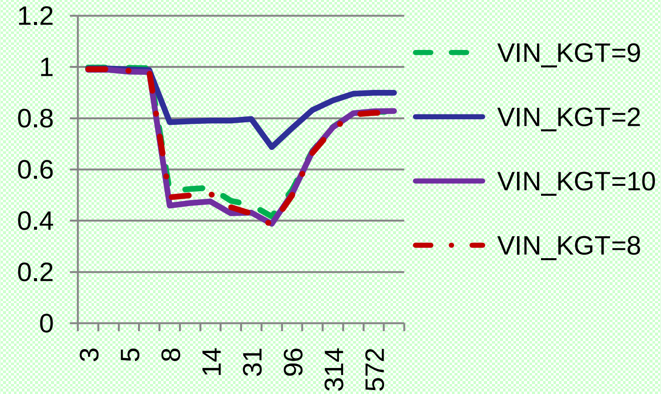
Groundtruth image, K = 12

Computing Experiment: Results

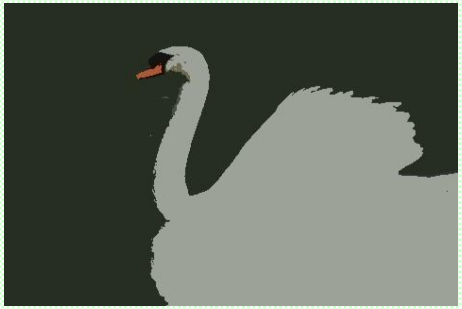




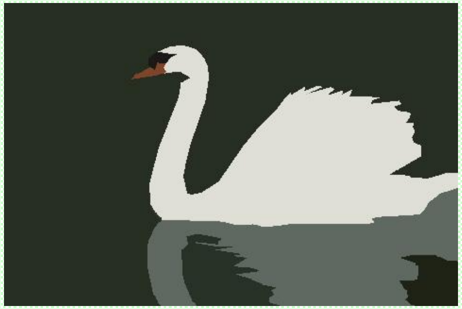




K=48

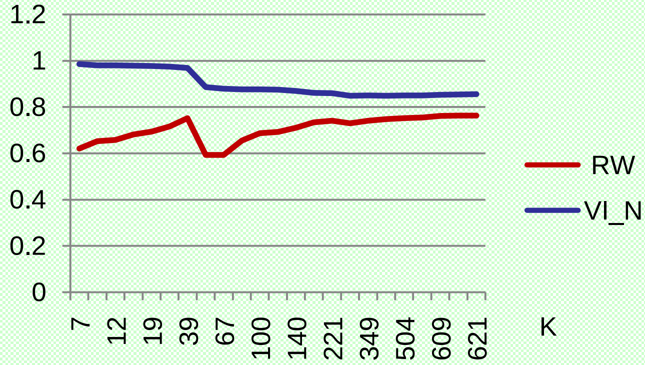


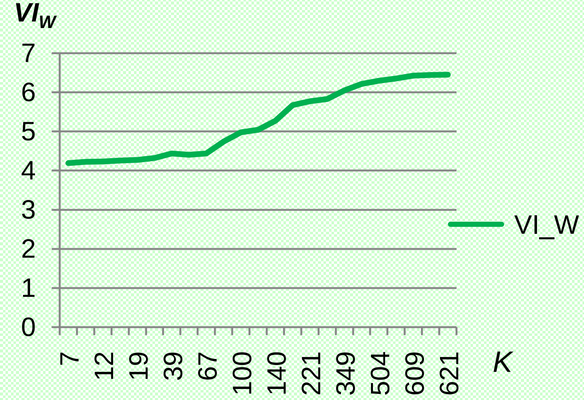
K=8

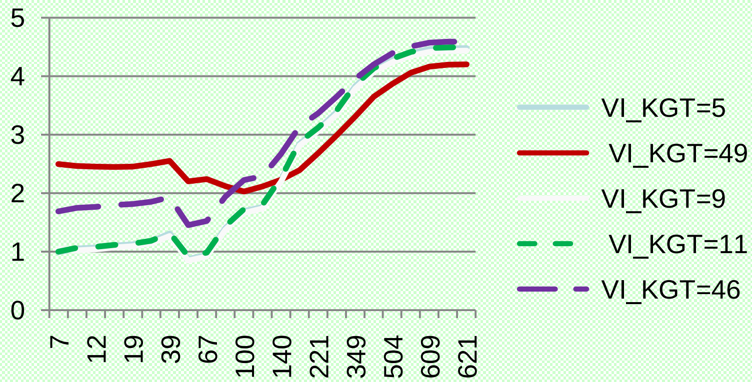


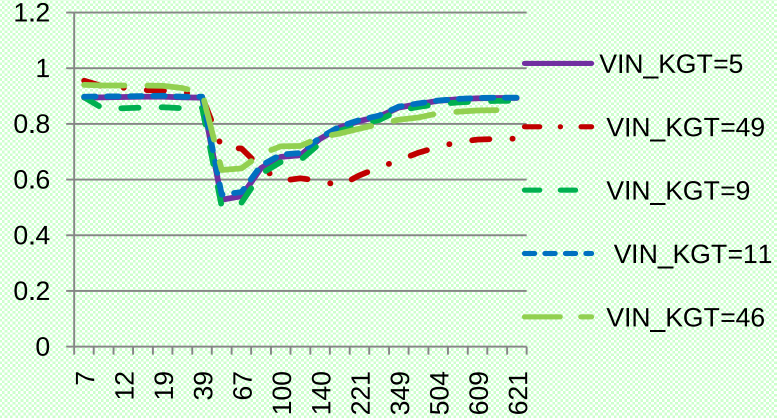
Groundtruth K=10

Computing Experiment: Results



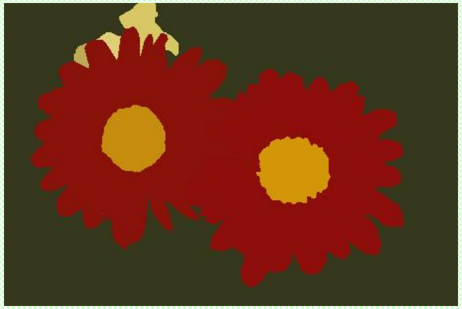






K=67

K=55



Groungtruth K=9

Groungtruth K=49

CONCLUSIONS

1. A technique for selecting the best segmentation from a set of images is developed.

2. Redundancy measure was proposed as a criterion of segmentation quality.

3. Computing experiment confirmed that the segmented image corresponding to a minimum of redundancy measure produces the suitable dissimilarity when compared with the original image.

4. The segmented image that was selected using the proposed criteria, gives the highest similarity to the groundtruth segmentations.

5. The future research will be aimed at the improving segmentation noise model and estimating the boundaries of application domain.