Hierarchical Partitions for Content Image Retrieval from Large-Scale Database

Dmitry Kinoshenko[1], Vladimir Mashtalir[1], Elena Yegorova[2], Vladimir Vinarsky[3]

[1]Kharkov National University of Radio Electronics, computer science faculty,
Lenin Ave., 14, 61166, Kharkov, Ukraine
{Kinoshenko, Mashtalir}@kture.kharkov.ua
[2]Wessex Institute of Technology,
Ashurst Lodge, Ashurst, Southampton, SO40 7AA, UK
Yegorova@kture.kharkov.ua
[3]Loewe Opta GmbH, Kompetenzzentrum Hannover,
Hildesheimer St. 140, D-30173 Hannover, Germany
vvinarski@loewe-komp.de

Abstract. Increasing of multimedia applications in commerce, biometrics, science, entertainments etc. leads to a great need of processing of digital visual content stored in very large databases. Many systems combine visual features and metadata analysis to solve the semantic gap between low-level visual features and high-level human concept, i.e. there arises a great interest in content-based image retrieval (CBIR) systems. As retrieval is computationally expensive, one of the most challenging moments in CBIR is minimizing of the retrieval process time. Widespread clustering techniques allow to group similar images in terms of their features proximity. The number of matches can be greatly reduced, but there is no guarantee that the global optimum solution is obtained. We propose a novel hierarchical clustering of image collections with objective function encompassing goals to number of matches at a search stage. Offered method enables construction of image retrieval systems with minimal query time.

1 Introduction

Short retrieval time independent of the database size is a fundamental requirement of any user friendly content-based image retrieval (CBIR) system. Visual contents of an image such as color, shape, texture, region relations play dominating role in propagation of features selection, indexing, user query and interaction, database management techniques. To search images in a large-scale image database traditionally queries 'ad exemplum' are used. Characteristics of different CBIR schemes, similarities or distances between the feature vectors of the query by example or sketch and those of the images collection are sufficiently full explored [1–7]. Essential efforts are devoted to synthesis and analysis of image content descriptors, namely color moments, histograms, coherence vectors, correlograms, invariant color features [8,9]; texture statistical and structural properties, determining by methods based on Fourier power spectra, Markov random fields, Gabor and wavelet transforms, fractal models, principal component analysis [10–13]; region-based and boundary-based features of shape, salient points in images [14–

16]; syntax and semantic representations [17,18]. They form feature spaces with reduced dimensions, however, when very large image collections are in use, matching methods are computationally expensive. One of the widespread approaches to minimize time outlay consists in application of various techniques of image database preliminary processing including clustering methods [7,10,18–20]. Thus a hierarchical mechanism of the query processing can be formed: one can seek suitable clusters in nested partitions with an arbitrary indexing scheme. This way the amount of matches can be greatly reduced, but traditional clustering methods do not guarantee the optimal result.

To optimize this CBIR scheme it is necessary to minimize a total number of matches at the retrieval stage. We propose a new hierarchical clustering which allows to construct images partitions into disjoint subsets so that firstly one can seek suitable class, then the most similar to the query subclass is chosen and so on. The exhaustive search is fulfilled only on the lower level of hierarchy. In the second section formalization of clustering problems is offered. In the third section clustering which guarantees the search for the minimal number of matching is discussed. Thus our contribution consists in development and theoretical ground of novel hierarchical partition construction for the fast content-based image retrieval in video databases.

2 Hierarchical partitions formalization for fast retrieval

In large databases an efficiency of image retrieval procedures with search 'ad exemplum' is determined in general by two characteristics: by reliability (in terms of precise identification of the required images) and computational complexity (generally, in sense of a matching operations amount) of algorithms used to pick up.

Usually first problem solution is coupled with the choice of features space which is adequate to application area. Image retrieval by example methods require that an index of proximity, or likeness, or affinity, or association has to be established between pairs of any images and such proximity functions can be expressed qualitatively and quantitatively, it can be represented by arbitrary metric or similarity (dissimilarity) measure. Further without loss of generality, let us suppose a metric is sufficient form of proximity function representation and we shall use notation $\rho(\circ,\circ)$, where arguments symbolize images in signal or feature space.

At second case, to avoid the combinatorial explosion a large collection of preliminary processing algorithms is available to reduce the search. Often clustering algorithms are used for computing complexity depreciation. Most of them are based on the two popular approaches: agglomerative hierarchical clustering and iterative partitional clustering. Hierarchical techniques organize data in a nested sequence of groups and partition techniques provide obtaining of one-level similar images classes.

Various heuristics are used to reduce a search complexity, but still there is no guarantee of optimal results. As a rule, to guarantee that a global optimum solution has been obtained, one has to examine all possible partitions, which is not computationally feasible. In this connexion let us consider original schemes with the best guaranteed results (with reference to a matches amount) in worst-case conditions.

Let $\mathrm{X} = \{x_v\}_{v\in\mathrm{V}}$ be a set of features representing images in the database (specifically it is possible to consider signal space also), hereinafter $\mathrm{V}, \Gamma, \mathrm{A}$ are index sets. Generally speaking, each feature vector belongs to Cartesian product $x_\alpha \in Z_1 \times Z_2 \times \ldots \times Z_r$ of arbitrary nature elements permitting to take into account properties of color distribution, shapes, texture, image semantic etc. Let $Y = \{y_\gamma\}_{\gamma \in \Gamma}$ be a feature set of an arbitrary query image.

The problem is in that under given query $y \in Y$ one needs to find the most similar image (or images) $x_v \in X$. In other words, it is necessary to provide $min_{v \in V}\, \rho(y, x_v)$ during minimum possible warranted time. If $Y \subseteq X$, the exact match retrieve is required. Such problems are typical for professional applications on industrial automation, biomedicine, crime prevention, medical diagnosis and prevention, social security, and other multi-spectral computer vision applications.

We shall name elements $[X]_\alpha$, $\alpha \in A$ of power set 2^X as clusters, if they correspond to the partition of set X. Let us consider such partitions that any elements of one cluster do not differ from each other more than on ε, i.e. $\forall\, x' \neq x''$ we have $[x'] = [x'']$, if $\rho(x', x'') \le \varepsilon$ and $[x'] \cap [x''] = \varnothing$ otherwise. The given or obtained value ε used at a clustering stage is connected with required accuracy of retrieval δ, if it is specified, as follows. There arise two cases:

$\delta > \varepsilon$ – any representative of the cluster nearest to the query y can be used as the image retrieval result, i.e. minimal number of matching operations is defined by the number of clusters; in other words it is necessary to provide

$$N_1 = card\{[X]_\alpha\} \to min; \qquad \textbf{(1)}$$

$\delta \le \varepsilon$ – the element of more detailed partition will be the result of the image retrieval. In simplest situations it is necessary to fulfill a single-stage clustering, i.e. to optimize retrieval under worst-case conditions we have to ensure

$$N_2 = card\{[X]_\alpha\} + max(card [X]_\alpha) \to min. \qquad \textbf{(2)}$$

At the multilevel clustering the repeated clusters search inside of already retrieved clusters is fulfilled and only on the last step required image is searched by complete enumeration. Let us assume that the cluster $[X^{(i-1)}]_p$ is selected on $(i-1)$ level of hierarchy from a condition $\rho(y, [X^{(i-1)}]_q) \to min$, $q = \overline{1, card\{[X^{(i-1)}]\}}$, i.e. $[X^{(i-1)}]_p = [X^{(i)}]_1 \cup \cup [X^{(i)}]_2 \cup \ldots \cup [X^{(i)}]_{\alpha_p}$, where for any k and l the equality $[X^{(i)}]_k \cap [X^{(i)}]_l = \varnothing$ holds. Then the minimization of matches amount is reduced to the clustering with the goal function

$$N_3 = \sum_{i=1}^{m-1} \{card\, [X^{(i)}]_{p,(i)} \big|_{x \in [X^{(i-1)}]_{p,(i-1)}}\} + max\,(card\, [X^{(m-1)}]_{p,(m-1)}) \to min\,, \qquad \textbf{(3)}$$

where m is a number of hierarchy levels, $[X^{(0)}]_{1,(0)} = X$. It should be noted, that a new metric to retrieve can be used on every stratum of nested clusters.

3 Hierarchical partitions with minimal matches amount

To find the multivalued map $\mathfrak{M}_\varepsilon : X \to \{[X]_\alpha\}$ providing the minimization of (3) let us introduce an iterative procedure $\mathcal{F}_k \big|_{\mathcal{H}_k(x)} : X \to 2^X$, $k = \overline{1, K}$, where

$$\mathcal{H}_k(x) = \mathcal{H}_{k-1}(x) \setminus \mathcal{F}_{k-1}(x),\ \mathcal{H}_0(x) = X,\ \mathcal{H}_{K+1}(x) = \varnothing,\ \mathcal{F}_0(x) = \varnothing,\ \mathcal{F}_k(x) = \bigcup_{i=1}^{l_k} x_i^* .$$

Here l_k is such that $\mathcal{H}_k^{l_k+1} = \varnothing$, where

$$\mathcal{H}_k^1 = \mathcal{H}_k, \quad \mathcal{H}_k^{i+1} = \{x_i^* \oplus \varepsilon\mathbb{B}^r\} \setminus \{x_i^*\}, \quad x_i^* = arg\ max_{x \in \mathcal{H}_k^i}\ card\{x_i \oplus \varepsilon\mathbb{B}^r\} .$$

Notation $\oplus$ designates the Minkowski sum, $\mathbb{B}^r$ is a unit ball in $\mathbb{R}^r$. The attracting points x_i^* actually give required multivalued map $\mathfrak{M}_\varepsilon$ in the form of

$$\mathfrak{M}_\varepsilon = \bigcup_{i=1}^{K} \mathcal{F}_k . \quad \textbf{(4)}$$

Recall, the value ε is the proximity measure within clusters. Also note that problem (2) is a special case of (3). We proved [21] that union (4) provides the solution of (1), i.e. the map $\mathfrak{M}_\varepsilon$ produces partitions (rather, maximal relative to inclusion posets) $\{[X]_\alpha\}$. Therefore these results can be exploited as initial clusters to get more detailed data partitioning. Finally minimizing (3) is reduced to the search of clusters cardinalities on each hierarchy level.

Let us suppose clusters $\{[X]_\alpha\}$ have cardinalities and multiplicities respectively $(M_1, s_1), \ldots, (M_t, s_t)$, $1 \le M_1 < M_2 < \ldots < M_t \le card\ X$, and the nesting hierarchy number is given (the optimum of hierarchy levels amount will be found further) and it equals m.

Objective function (3) can be rewritten $\forall \tau_1 \in [M_{n-1}, M_n] \cap \mathbb{N}$ ($M_0 = 1$) as

$$f(\tau_1, \ldots, \tau_{m-1}) = \sum_{j=1}^{n-1} s_j + \sum_{j=n}^{t} \lceil M_j / s_j \rceil s_j + \lceil \tau_1 / \tau_2 \rceil + \ldots + \lceil \tau_{m-2} / \tau_{m-1} \rceil + \tau_{m-1}, \quad \textbf{(5)}$$

where values $\tau_1 > \tau_2 > \ldots > \tau_m = 1$ correspond to coefficients of sequential clusters partitioning, $\lceil \circ \rceil$ denotes a ceiling function.

Points $\tau_1^*, \tau_2^*, \ldots, \tau_{m-1}^*$, at which the global minimum of function $f(\tau_1, \tau_2, \ldots, \tau_{m-1})$ is obtained on the set $[1, M_t] \cap \mathbb{N}$, represent required parameters for hierarchical partitions. Local minimum of this function we shall search in the partial segments $[M_{n-1}, M_n]$ $(n = \overline{1,t})$ and global minimum we shall get by search among the obtained values.

For discrete function minimum search we shall first carry out continuous minority function minimizing

$$\varphi(u_1, u_2, \ldots, u_{m-1}) = \sum_{j=n}^{t} M_j s_j / u_1 + \sum_{j=1}^{n-1} s_j + u_1/u_2 + \ldots + u_{m-3}/u_{m-2} + u_{m-1} . \quad \textbf{(6)}$$

Errors of transfer to minority functions are defined by the expression

$$f(u_1, u_2, \ldots, u_{m-1}) - \varphi(u_1, u_2, \ldots, u_{m-1}) \le \tilde{\Delta}_n = \sum_{j=1}^{n-1} s_j + m - 2 . \quad \textbf{(7)}$$

Consider the function (6) minimizing problem for $u_1 \in [M_{n-1}, M_n]$. Let us emphasize that it has additive form, all items are positive and each item of type u_j/u_{j+1} $(j = \overline{2, m-2})$, evidently, sequentially depends on two variables only. Taking into account these properties, first we can fix $u_1, u_2, \ldots, u_{m-2}$, then we have to find

$$u_{m-2}/u_{m-1} + u_{m-1} \to \min_{u_{m-1}} ,$$

whence it follows that $u_{m-1}^* = \sqrt{u_{m-2}}$. Substituting this value into (7) we get

$$\varphi(u_1, u_2, \ldots, u_{k-2}) = \sum_{j=n}^{t} M_j s_j / u_1 + \sum_{j=1}^{n-1} s_j + u_1/u_2 + \ldots + u_{m-3}/u_{m-2} + 2\sqrt{u_{m-2}} .$$

Let us find then

$$u_{m-3}/u_{m-2}+\sqrt{u_{m-2}} \to \min_{u_{m-2}},$$

i.e. $u_{m-2}^{*}=u_{m-3}^{2/3}$. Continuing this process we come to the relations

$$u_m^*=1,\ u_{m-1}^*=(u_{m-2}^*)^{1/2},\ \dots,\ u_2^*=(u_1)^{(m-2)/(m-1)}. \quad \mathbf{(8)}$$

Thus we finally arrive at

$$\varphi(u_1)=\varphi(u_1,u_2^*,u_3^*,\dots,u_{k-1}^*)=\mathrm{A}_n/u_1+(k-1)u_1^{1/(k-1)}+\mathrm{B}_{n-1},$$

where $\mathrm{A}_n=\sum_{j=n}^{t}\mathrm{M}_j s_j$, $\mathrm{B}_{n-1}=\sum_{j=1}^{n-1}s_j$.

Let us analyze this function. First notice that $\varphi(u_1)$ is an unimodal function. Indeed, $\varphi'(u_1)=-\mathrm{A}_n/u_1^2+1/u_1^{(m-2)/(m-1)}=0$, therefore, $u_1^*=\left(\sum_{j=n}^{t}\mathrm{M}_j s_j\right)^{m/(m-1)}$. Further since $sign\,\varphi'(u_1)=sign\,(-\mathrm{A}_n+u_1^{m/(m-1)})$, function $\varphi(u_1)$ decreases on $]0,u_1^*[$ and increases on $]u_1^*,\infty[$. Consequently, minorant (6) reaches its minimum value either at the point u_1^* if $u_1^*\in[\mathrm{M}_{n-1},\mathrm{M}_n]$, or at boundary points of this interval. Thereby,

$$\min_{u_1\in[\mathrm{M}_{n-1},\mathrm{M}_n]}\varphi(u_1)=\min\{\varphi(\mathrm{M}_{n-1}),\varphi(\mathrm{M}_n),\varphi(\max(\min(u_1^*,\mathrm{M}_n),\mathrm{M}_{n-1}))\}.$$

Let u_1^{**} be a point at which the minimum is achieved

$$u_1^{**}=\arg\min_{[\mathrm{M}_{n-1},\mathrm{M}_n]}\varphi(u_1)\in\{\mathrm{M}_{n-1},\mathrm{M}_n,\ \max(\min(u_1^*,\mathrm{M}_n),\mathrm{M}_{n-1})\},$$

then from (8) we get

$$u_2^*=(u_1^{**})^{(m-2)/(m-1)};\ u_3^*=(u_1^{**})^{(m-3)/(m-1)},\ \dots,\ u_{m-1}^*=(u_1^{**})^{1/(m-1)},\ u_m^*=1.$$

Let us find the ranges of definition of $u_1,u_2,\dots,u_{m-1}$ at the return to the discrete goal function. Denote $\psi_i(u_1)=(u_1)^{(m-i)/(m-1)}$ then obviously $\varphi(u_1,u_2,\dots,u_{m-1})\ge$ $\ge\varphi(u_1,\psi_2(u_1),\psi_3(u_1),\dots,\psi_{m-1}(u_1))$. In the result of minorant minimization we get $\varphi(u_1^*,u_2^*,\dots,u_{m-1}^*)=\mathrm{B}^*$. If we choose some value $0\le\Delta\le\tilde{\Delta}_n$ from (7) we have to solve inequality $\varphi(u_1,u_2,\dots,u_{m-1})\le\mathrm{B}^*+\Delta$. Denote a solution set as $\mathrm{P}(\Delta)$. It is easy to prove that $\mathrm{P}(\Delta)\subset\mathrm{P}_1(\Delta)\times\psi_2(\mathrm{P}_1(\Delta))\times\dots\times\psi_{m-1}(\mathrm{P}_1(\Delta))$.

In that way, finding variables changing intervals $\tau_1,\tau_2,\dots,\tau_{m-1}$ should be started from search of $\mathrm{P}_1(\Delta)$. From (7) we find

$$u_1\left(\mathrm{B}^*+\Delta-\sum_{j=1}^{n-1}s_j-(m-1)u_1^{1/(m-1)}\right)\ge\sum_{j=n}^{t}\mathrm{M}_j s_j. \quad \mathbf{(9)}$$

Under $m\ge4$ inequality (9) can be solved only with numerical methods. Taking into consideration integer-valued character of the search $\tau_1,\tau_2,\dots,\tau_{m-1}$ and relation (7) we shall find the interval $\Delta_1=[\tau_1',\tau_1'']\cap\mathbb{N}$, for example, by dichotomy of intervals $[\mathrm{M}_{n-1},\lceil u_1^*\rceil]\cap\mathbb{N}$ and $[\lfloor u_1^*\rfloor,\mathrm{M}_n]\cap\mathbb{N}$, ($\lfloor\circ\rfloor$ denotes a floor function) or $[\mathrm{M}_{n-1},\mathrm{M}_n]\cap$ $\cap\mathbb{N}$ if $u_1^*\in\{\mathrm{M}_{n-1},\mathrm{M}_n\}$. So we finally get

$$\forall j\in\{2,3,\dots,m-1\}\quad \Delta_j=[\lfloor\psi_m(\tau_1')\rfloor,\lceil\psi_m(\tau_1'')\rceil]\cap\mathbb{N}.$$

It follows from the above that minimization of retrieval operations number

$$f(\tau_1,\tau_2,\dots,\tau_{m-1})\to\min_{[\mathrm{M}_{n-1},\mathrm{M}_n]} \quad \mathbf{(10)}$$

can be done on the intervals $\tau_1\in\Delta_1\cap\mathbb{N},\ \tau_2\in\Delta_2\cap\mathbb{N},\dots,\ \tau_{m-1}\in\Delta_{m-1}\cap\mathbb{N}$.

Let us introduce the hierarchical clusters τ_i -decomposition cortege concept: it is an arbitrary set of integer values $\{\tilde{\tau}_i, \tau_{i+1}, \ldots, \tau_{m-1}\}$ satisfying conditions

$$\forall i \in \{1, 2, \ldots, m-1\} \Rightarrow \tau_i \in \Delta_i \bigcap \mathbb{N}\,,\ \forall i, j \in \{1, 2, \ldots, m-1\} : j \geq i \Rightarrow \tau_j > \tau_{j+1}\,.$$

Thus, the problem is reduced to the search among all hierarchical clusters τ_i-decomposition vectors such, that the requirement (3) (or (10) what is the same) is met. To solve this problem we shall first draw a backward recurrence of function (5).

The hierarchical clusters τ_i -decomposition cortege we name optimal if under fixed value $\tilde{\tau}_i$ the function

$$F(\tau_{i+1}, \ldots, \tau_{m-1}) = \lceil \tilde{\tau}_i / \tau_{i+1} \rceil + \lceil \tau_{i+1} / \tau_{i+2} \rceil + \ldots + \lceil \tau_{m-2} / \tau_{m-1} \rceil + \tau_{m-1} \qquad \mathbf{(11)}$$

has a global optimum on the considered set $(\tilde{\tau}_i, \tau_{i+1}, \ldots, \tau_{m-1})$.

On the base of dynamic programming paradigm it is easy to show that the hierarchical clusters τ_j -decomposition cortege belonging to the τ_i -decomposition optimal cortege ($j < i$) is also optimal. Thus we can formulate hierarchical clustering method on the base of backward recurrence, i.e. starting with parameter of τ_{m-1} -decomposition and finishing with τ_1 -decomposition cortege. At each step the concatenation of τ_i -decomposition optimal cortege with an element of higher hierarchical level is carried out. Let us consider this procedure closely.

Under given number of hierarchy levels the input data is a set of integer intervals $\{\Delta_1, \Delta_2, \ldots, \Delta_{m-1}\}$. Starting from the lowest level of hierarchy we assume that all the parameters of τ_{m-1} -decomposition are optimal. On the same level we form an initial set $T_{m-1} = \{\tau_{m-1}\}_{\tau_{m-1} \in \Delta_{m-1}}$ of potential optimal decomposition corteges. Then on the $(m-2)$ level we consider sets $\{\tau_{m-2}, \tau_{m-1}\}$, choosing such $\{\tilde{\tau}_{m-2}, \tau_{m-1}\}$ which provides the minimum of function (11). When we find optimal τ_{m-2} -decomposition corteges, we shall modify the set of potential optimal decomposition corteges $T_{m-2} = \{\tilde{\tau}_{m-2}, \tau_{m-1}\}_{\substack{\tilde{\tau}_{m-2} \in \Delta_{m-2} \\ \tau_{m-1} \in T_{m-1}}}$. Continuing this procedure we obtain

$$\{\tilde{\tau}_i^j,\ \tau_{i+1}, \ldots,\ \tau_{m-1}\} = arg \min_{\{\tau_{i+1}, \tau_{i+2}, \ldots, \tau_{m-1}\} \in T_{i+1}} \{\tau_i^j \,||\, \{\tau_{i+1},\ \tau_{i+2}, \ldots,\ \tau_{m-1}\}\}\,,$$

where $||$ denotes a concatenation operation; $T_i = \{\tilde{\tau}_i^j, \tau_{i+1}, \ldots, \tau_{m-1}\}$.

The selection of the optimal cortege $\{\tau_1^*, \tau_2^*, \ldots, \tau_{m-1}^*\}$ i.e. the set of function (10) arguments is carried out on the last step of the backward recurrence by searching

$$\{\tau_1^*, \tau_2^*, \ldots, \tau_{m-1}^*\} = arg \min_{\{\tau_1, \tau_2, \ldots, \tau_{m-1}\} \in T_1} f(\tau_1, \tau_2, \ldots, \tau_{m-1})\,, \qquad \mathbf{(12)}$$

where the number of operations is within the value $card\,[\Delta_1 \bigcap \mathbb{N}]$.

Thus problem (3) solution is obtained in the form (12) on the partial interval $[M_{n-1}, M_n]$ $(n = \overline{1,\ t};\ M_0 = 1)$. The global optimum on the whole domain of clustering parameters $[M_1, M_t]$ is chosen among the partial records. So, we have found parameters of clusters partitions at each level of hierarchy. Now there is a need only to divide clusters separately one from other.

It should be emphasized that offered approach enables to use one-parameter sequential optimization on the Cartesian product of multivalued maps $\mathfrak{M} : X \to \{[X]_\alpha\}$ preimages to search the initial (maximal relative to inclusions) clusters, the degree of objects similarity in these clusters and the stratification coefficient. It can be explained by an independence of the indicated parameters and an absence of principal restrictions to time outlays at preliminary processing of data in CBIR systems.

4 Results and outlook

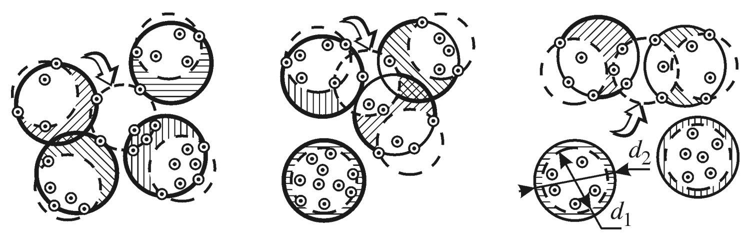
Till now we considered clustering with given parameters, namely at known maximum diameter of clusters at solution (1) and amount of strata at solution (3). It is clear that for the reliability growth it is expedient to use sufficiently small values ε, but then a number of matches is increased. Matches number decreasing requires optimization of clusters powers, that is reached by increase of ε. In other words, at $\varepsilon \to a$, where

$$a = min\,\{\lambda \geq 0 : (x' \oplus \lambda\mathbb{B}^r) \bigcap (x'' \oplus \lambda\mathbb{B}^r) \neq \varnothing \;\forall\, x', x'' \in \mathrm{X}\},$$

the problem is reduced to the exhaustive search. At the same time at $\varepsilon \to b$, where

$$b = min\,\{\lambda \geq 0 : \mathrm{X} \subseteq x \oplus \lambda\mathbb{B}^r \;\; \forall x \in \mathrm{X}\},$$

the number of matches also tends to $card\ \mathrm{X}$. The conflict between two criteria (combinatorial complexity of matches and its reliability) is eliminated for each images configuration at a stage of preliminary processing when there are no key constraints to time outlay. Furthermore to reach desired degree of accuracy and reliability, it is necessary to solve a multiextremal problem. Indeed, fig. 1 illustrates changes of ε from d_1 to d_2 for problem (2): three cases of clusters amount decreasing are shown (immersed sets are indicated by arrows), but matches number can be increased (*a*), fixed (*b*), decreased (*c*).



a *b* *c*

Fig. 1. Clusters merging examples under ε changing

Hence the questions concerning rational choice of clustering parameters with respect to CBIR effectiveness are of great significance. A search of optimal ε is a subject of further inquiry. Here let us find the optimal number of hierarchy levels. By analogy with (6), let us introduce minority function for $[1, \mathrm{M}_t]$

$$\varphi_g(u_1, u_2, \ldots, u_{m-1}) = \textstyle\sum_{j=n}^{t} \mathrm{M}_j s_j / u_1 + u_1/u_2 + \ldots + u_{m-3}/u_{m-2} + u_{m-1}\,.$$

Then from (7) we have $\varphi(u_1, u_2, \ldots, u_{m-1}) = m(\sum_{j=n}^{t} \mathrm{M}_j s_j\,)^{-m} \to min_{m \in \mathbb{N}}$, i.e.

$$m \in \{\lfloor \mathrm{L} \rfloor, \lceil \mathrm{L} \rceil\}, \quad \mathrm{L} = ln\,(\textstyle\sum_{j=1}^{t} \mathrm{M}_j s_j)\,.$$

Thus to find optimal number of hierarchy levels it is sufficiently to check only these two values.

The experimental researches of image clustering were carried out in a 29-dimensional feature space. Five groups of features characterizing properties of gray-level functions, their planar peculiarities, objects shapes, invariance to one-parameter geometric transformations and also to Euclidean group of similarities were chosen [21].

As a database a set of images of 108 cars was selected (see fig. 2). Different acquisition conditions were modeled for each image, namely, additive Gaussian and uniform

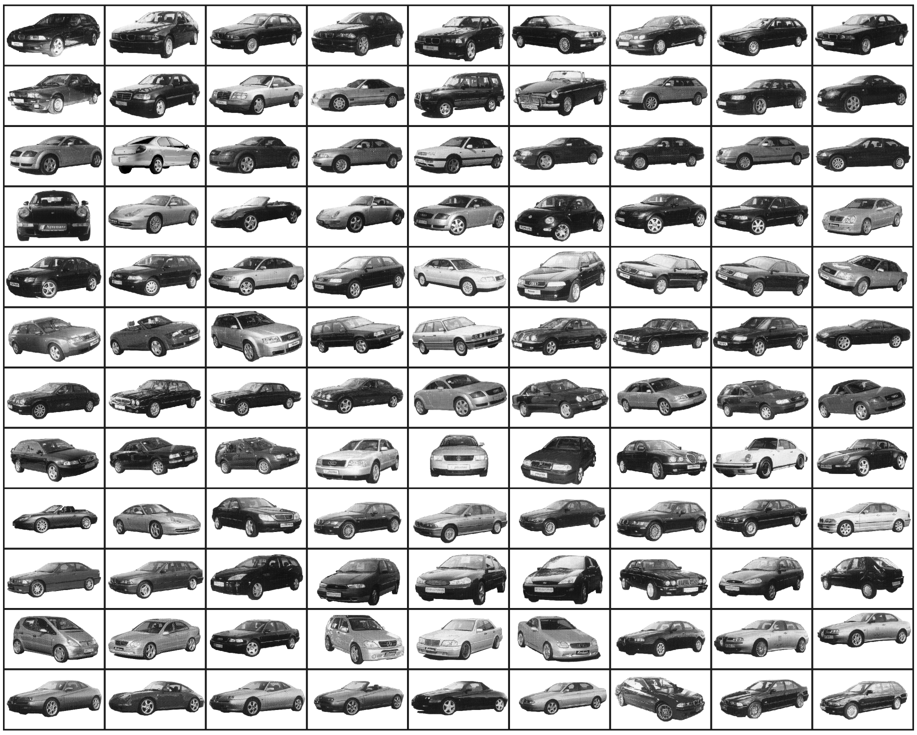


Fig. 2. Basic set of images

noise, adding the point sources of lighting, discordance of horizontal and vertical scanning, image blurring, image motion modeling, brightness, contrast, lighting intensity variations, linear histogram transformations (see fig. 3). As a result the base set of images was expanded up to 1080 frames. Besides for the image analysis conditioned by variations of a mutual location and/or orientation of object and videosensor, groups of geometric transformations such as rotation, skewing, scaling, perspective distortions, Euclidean similarities (shift, scaling and rotation are acting simultaneously) were used. Thus, the number of the images varied from 1080 up to 6480.

Before clustering all features were normalized to segment [0,1] then they were attributed by weight coefficients. Mentioned above feature families were used either in total, or separately, or selective by elements. As an example, in fig. 4 the fragment of retrieval via (3) with Hausdorff metric is shown. The analysis of results allows to draw the conclusion that at the correct features selection (namely choice of necessary and sufficient feature set, which has adequate information efficiency) under conditions mentioned

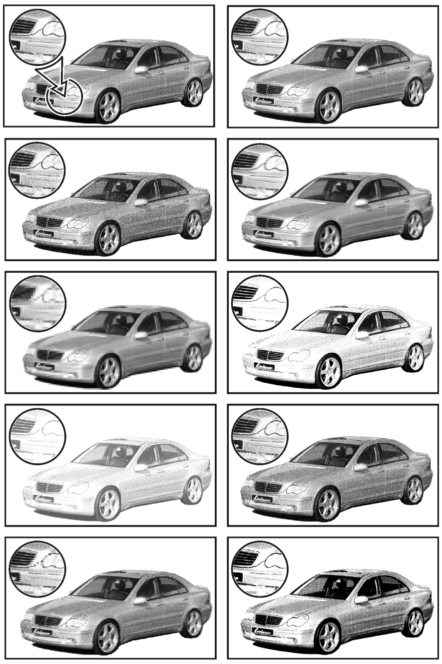


Fig. 3. Variants of image acquisition

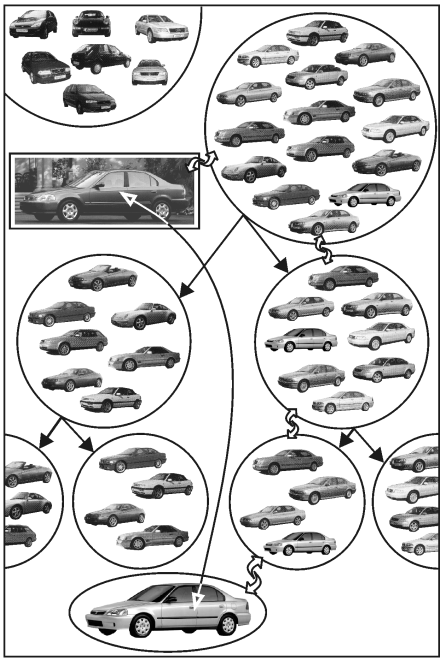


Fig. 4. Fragment of image retrieval

number of query matching reduces in 25–110 times as compared to the exhaustive search virtually with idem reliability. In comparison with the most popular hierarchical clustering techniques (nearest and furthest neighbor, median and centroid algorithms, Ward's minimum variance) matches number is diminished by factor of $N = 6\ldots17$. It should be emphasized that our approach guarantees minimal number of matches for arbitrary features configuration while traditional clustering methods do it on average. From the practical standpoint the solution of problem (4) often is the most convenient by the virtue of large-scale database processing simplicity. We have ascertained that under minimization (2) the number of matches aspires to $\lambda\sqrt{card\ \mathrm{X}}$ at essential growth $card\ \mathrm{X}$ (in ideal case $\lambda = 2$).

To provide fast access to an image in database with queries 'ad exemplum' a novel method of hierarchical clustering has been offered and investigated. Proposed method ensures minimal number of matches at CBIR stage for arbitrary images collection in feature or signal space.

In addition we shall indicate that the offered stratified clustering enables to take into account a minimum of clusters diameters, maxima of intercluster or interlevel distances. Moreover we have proposed the clustering method guaranteeing the best result in a worst-case condition, i.e. often spared hardware-software resource can be used with the purpose of reliability CBIR systems increase. Obeying a key criterion (minimum combinatorial complexity of CBIR), it is also possible to search desired nested clusters by modifications of well-known clustering techniques keeping their advantages.

References

1. Chen, Y., Wang, J.Z.: Image categorization by learning and reasoning with regions. Journal of Machine Learning Research, Vol. 5 (2004) 913-939

2. Li, J., Wang, J.Z.: Automatic linguistic indexing of pictures by a statistical modeling approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25 (2003) 1075-1088

3 Müller, H., Müller, W., Squire, D.McG., Marchand-Maillet, S., Pun, T.: Performance Evaluation in Content-Based Image Retrieval: Overview and Proposals. Pattern Recognition Letters, Vol. 22 (2001) 593-601

4. Veltkamp, R., Burkhardt, H., Kriegel, H.-P. State-of-the-Art in Content-Based Image and Video Retrieval. Kluwer Academic Publishers (2001)

5. Greenspan, H., Dvir., G., Rubner, Y.: Context-Dependent Segmentation and Matching in Image Database. Computer Vision and Image understanding, Vol. 93 (2004) 86-109

6. Dimai, A.: Assessment of Effectiveness of Content-Based Image Retrieval Systems. Lecture Notes in Computer Science, Springer-Verlag, Berlin-Heidelberg, Vol. 1614 (1999) 525-532

7. Rui, Y., Huang, T., Chang, S.: Image Retrieval: Current Techniques, Promising Directions and Open Issues. Journal of Visual Communication and Image Representation, Vol.10 (1999) 39-62

8. Yanai, K., Shindo, M., Noshita, K.: A Fast Image-Gathering System from the World-Wide Web Using a PC Cluster. Image and Vision Computing, Vol. 22 (2004) 59-71

9. Bauckhage, C.,Braun, E., Sagerer, G.: From Image Features to Symbols and Vice Versa – Using Graphs to Loop Data – and Model-Driven Processing in Visual Assembly Recognition. International Journal of Pattern Recognition and Artificial Intelligence, Vol.18 (2004) 497-517

10. Manjunath, B.S., Ma, W.Y.: Texture Features for Browsing and Retrieval of Large Image Data. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18 (1996) 837-842

11. Celebi, E., Alpkocak, A.: Clustering of Texture Features for Content Based Image Retrieval. Lecture Notes in Computer Science, Springer-Verlag, Heidelberg, Vol. 1909 (2000) 216-225

12. Peng, J., Bhanu, B., Qing, S.: Probabilistic Feature Relevance Learning for Content-Based Image Retrieval. Computer Vision and Image Understanding, Vol. 75, (1999) 150-164

13. Cox, I.J., Miller, M.L., Minka, T.P., Papathomas, T., Yianilos, P.N.: The Bayesian Image Retrieval System, PicHunter: Theory, Implementation, and Psychophysical Experiments. IEEE Transactions on Image Processing, Vol. 9 (2000) 20-37

14. Carson, C., Belongie, S., Greenspan, H., Malik, J.: Region-Based Image Querying. Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries (CVPR' 97) (1997) 42-49

15. Tian, Q., Sebe, N., Lew, M.S., Loupias, E., Huang, T.S.: Image Retrieval Using Wavelet-Based Salient Points. Journal of Electronic Imaging, Vol. 10 (2001) 835-849

16. Cinque., L., De Rosa, F., Lecca, F., Levialdi, S.: Image Retrieval Using Resegmentation Driven by Query Rectangles. Image and Vision Computing, Vol. 22 (2004) 15-22

17. Santini, S., Gupta, A., Jain, R.: Emergent Semantics through Interaction in Image Databases. Knowledge and Data Engineering, Vol. 13 (2001) 337-351

18. Sheikholeslami, G., Chang, W., Zhang, A.: SemQuery: Semantic Clustering and Querying on Heterogeneous Features for Visual Data. IEEE Transactions on Knowledge and Data Engineering, Vol. 14 (2002) 988-1002

19. Käster, T., Wendt, V., Sagerer, G.: Comparing Clustering Methods for Database Categorization in Image Retrieval. Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg, Vol. 2781 (2003) 228-235

20. Mashtalir, V.P., Yakovlev, S.V.: Point-Set Methods of Clusterization of Standard Information. Cybernetics and Systems Analysis, Kluwer Academic Publishers, Vol. 37 (2001) 295-307

21. Kinoshenko, D., Mashtalir, V., Orlov, A., Yegorova, E.: Method of Creating of Functional Invariants under One-Parameter Geometric Image Transformations. Lecture Notes in Computer Science, Springer-Verlag, Berlin-Heidelberg, Vol. 2781 (2003) 68-75