

**VARNA FREE UNIVERSITY "CHERNORIZETS HRABAR"  
INSTITUTE OF MATHEMATICS AND INFORMATICS OF THE  
BULGARIAN ACADEMY OF SCIENCE**

**INFORMATICS  
IN THE SCIENTIFIC KNOWLEDGE**

**INTERNATIONAL SCIENTIFIC CONFERENCE**

**2010**



# HIGH LEVEL SEMANTIC RETRIEVAL IN WEB BASED ART PAINTING DIGITAL REPOSITORIES

Peter Stanchev, Phillip Ein-Dor, Krassimira Ivanova, Iliya Mitov

**Abstract:** *In this paper, we present, compare, and improve existing approaches to the problem of high level semantic retrieval in art painting digital repositories. A special prototype software system for studying the image analysis and retrieval is also presented.*

**Keywords:** *High level semantic retrieval, art painting, art repositories*

## 1 Introduction

Current search engines headed by Google are in the centre of our information age. Google answers daily more than 200 million queries against over 30 billion items. However, the search power of these engines is typically limited to text and its similarity. Since less than 1% of the Web data is in textual form, the rest being of multimedia/streaming nature, we need to extend our next-generation search to accommodate these heterogeneous media. The new search machines must combine search according to textual information or other attributes associated with the files with abilities of extracting information from the content, which is scope of action of Content-Based Image Retrieval (CBIR). However, the typical database size is in the order of thousands of images. Very recent publicly-available systems, such as ImBrowse2, Tiltomo3 or Alipr4, claim to index hundreds of thousands of images. There is a high discrepancy between these numbers and the volumes of images available on the current Web.

In this context, we developed an architecture for indexing and searching in picture collections according to visual characteristics of their content. This source has also the advantage of having associated user-defined textual information, which could be used for future experiments with combining search on text and visual content. CBIR systems should leap towards Web-scale datasets. In this paper, we utilize our experience to build principles of a system, with the corresponding descriptive features, to be used in experimenting with new techniques for similarity searching. We have to tackle the non-trivial process of image crawling and descriptive feature extraction (using some of the MPEG-7 features). Some of the characteristics of the web visual semantics were outlined during the research. As an experimental base we use art painting digital repositories. Especially for this article, images from Israel artists, presented in Givon Art Gallery ([www.qivonartgallery.com](http://www.qivonartgallery.com)), are used.

The rest of the paper is organized in the following way. Section 2 is a review of existing art digital repositories. In Section 3 is a brief overview of MPEG 7 descriptors used. Section 4 presents proposed algorithm and program realization. Section 5 describes the experiments made by contemporary Israel artists, presented in the Web. Finally conclusions and some future trends are outlined.

## **2 Art Painting Digital Repositories**

Cultural heritage objects are rich in content, describing events, monuments, places, people etc. and distributed across different locations. The users can formulate queries using different modalities such as free text, similarity matching, or metadata. The cultural image digital repositories can employ similar techniques as in generic systems in order to solve standard queries for searching objects. The specifics of observed objects profile some additional tasks, which are interesting to analyze. In the area of art paintings retrieval, sensitive, semantic and aesthetic gaps are quite a real problem.

In 2004 David Mattison, named as master of the online archive universe, published a series of lectures in Searcher magazine focusing on the state-of-the-art of available Web resources and image databases, current techniques for image retrieval, and finishing with mention of national collections that document the art history of Western civilization from medieval times through the 19th century [Mattison, 2004]. The creators of this wealth of image databases, art collections and guides usually are academic, librarian, commercial, and private art museums and galleries, amateur and professional art historians, artist sites, commercial image agencies, auction houses (usually on a temporary basis), etc.

Several different applications in the field of Fine Arts have led to specialized digital picture processing. During the years, many successful projects in the field of processing very large high quality color images were funded by Esprit, Impact, Raphael, and IST programs. They provide the most conventional image processing for the museum, such as geometric correction, registration, mosaicing, etc. [Maitre et al, 2001]. The European Union has funded numerous digital culture research and development initiatives. The EU's CORDIS (Community Research & Development Information Service) (<http://www.cordis.lu/>) is the primary resource for learning about past and current R&D projects involving art.

Some of the projects, such as Vasari (1989-1992) and Marc (1995-1996), are focused on digital acquisition, storage and handling of colorimetric high-definition images of paintings (up to 2GB per image) for different galleries and museums in Western Europe. Project Crisatel

(2001-2004) developed equipment for the direct fast capture of paintings, with a new ultra-high definition multi-spectral scanner in order to make spectrometric analysis of varnish layers to allow the effect of an aged varnish to be subtracted from an image of a painting. Project FingArtPrint (2005-2008) is aimed at combining 3D surface scanning and multispectral imaging in order to create a unique data record of the object which can be compared to check its authenticity.

Other projects and initiatives are aimed at establishing image repositories. One of the first projects in this direction was NARCISSE (1990-1992), which has created a very high digitized image bank, supervised by a multilingual text database (in German, French, Italian and Portuguese). The objective of the project Artiste (2000-2002) was to develop and prove the value of an integrated art analysis and navigation environment aimed at supporting the work of professional users in the fine arts. The environment has exploited advanced image content analysis techniques, distributed hyperlink-based navigation methods, and object relational database technologies, using existing metadata standards and indexing schemes. Artiste has integrated art collections virtually while allowing the owners of each collection to maintain ownership and control of their data, using the concept of distributed linking.

The latest and biggest project in this direction is Europeana (<http://www.europeana.eu/portal/>), funded by the European Commission and the member states. The idea of Europeana was born in 2005, when the European Commission announced its strategy to promote and support the creation of a European digital library, as a strategic goal within the European Information Society i2010 Initiative, which aims to foster growth and jobs in the information society and media industries. The goal is to make European information resources easier to use in an online environment. It will build on Europe's rich heritage, combining multicultural and multilingual environments with technological advances and new business models. Europeana.eu went live on 20 November 2008 and till now more than 6 millions digital items (Images – paintings, drawings, maps, photos and pictures of museum objects; Texts – books, newspapers, letters, diaries and archival papers; Sounds – music and spoken word from cylinders, tapes, discs and radio broadcasts; Videos – films, newsreels and TV broadcasts) are available.

Some of very successful worldwide digital repositories are created independently; as a result of collaboration between different state or public institutions; or by supporting concrete company. Usually the money for establishing and supporting such initiatives are received by donations; by sparing out of the budgets of the institutions; or by

allocating funds from the profit of the companies gained with other activities (for instance Artchive is supported by art.com, which is company for selling digital posters of fine art and decor items).

Some examples of art painting repositories are Artcyclopedia (<http://www.artcyclopedia.com/>), Artchive (<http://www.artchive.com/>), OCAIW (Orazio Centaro's Art Images on the Web) (<http://www.ocaiw.com/>), WebMuseum (<http://www.ibiblio.org/wm/>), Web Gallery of Art (<http://www.wga.hu/>), Backtoclassics (<http://www.Backtoclassics.com/>), Olga's Gallery (<http://www.abcgallery.com/>).

Currently a lot of museums have online galleries, supplying access to the museum's collections. The search engines of some of them are based only on metadata information using categories such as artist, title of work, subject, chronology and reference number (for instance <http://www.museodelprado.es/en/the-collection/online-gallery/>), in other sites attempts to implement CBIR techniques during the search process are included (such as <http://www.hermitagemuseum.org>).

Let's mention also another type of site, which supplies access to the sites of the museums and online galleries. For instance Museumland (<http://www.museumland.com/>) is the worldwide portal to museums and cultural heritage with more than 11 000 links.

### 3 MPEG 7 Descriptors

MPEG 7 descriptors are complex descriptors, which provide a good presentation of different types of visual features. The description of the structure of MPEG 7 descriptors and algorithms is given in [MPEG, 2002]. These complex structures need specific processing and cannot be put directly into generic classification algorithms.

Here we give a brief explanation of each MPEG 7 descriptor examined in our work and which values we use in further processing:

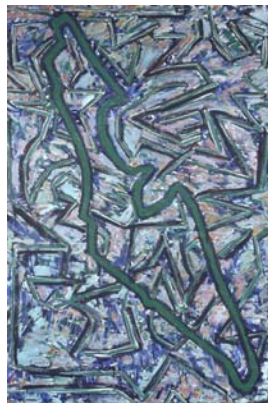
- Scalable Color (SC) represents the color histogram in the HSV color space, encoded by a Haar transform. For presenting the image or a selected part, Scalable Color needs a vector with 64 attributes;
- Color Layout (CL) specifies the spatial distribution of colors using YCbCr color space. We use the first quantized DCT coefficient of the Y, Cb and Cr components, the next five successive quantized DCT coefficients of the Y component and the next two quantized DCT coefficients of the Cb and Cr component. Thus, the Color Layout vector has 12 attributes;
- Color Structure (CS), which specifies both color content and the structure of the content. The descriptor expresses local color structure in an image by means of a structuring element that is

composed of several image samples. We use a vector with 64 attributes for representing the Color Structure;

- Dominant Color (DC). We reconfigured the presentation of this descriptor as three vectors, representing distribution of quantized hue, saturation and luminance. Such method is already precisely described and used in previous work on APICAS in [Ivanova and Stanchev, 2009]. After this quantization we receive a vector with 23 attributes (13 for hue + 5 for saturation + 5 for luminance);
- Edge Histogram (EH) specifies the spatial distribution of five types of edges in local image regions (four directional edges – vertical, horizontal, 45 degree, 135 degree and one non-directional). Edge Histogram descriptor produces a vector with 80 attributes;
- Homogeneous Texture (HT) characterizes the region texture using the energy and energy deviation in a set of frequency channels. A vector with 60 attributes is used, presenting Energy and Energy Deviation.

Some of the MPEG 7 descriptors are alternative. For instance, Scalable Color, Color Layout, Color Structure and Dominant Color concern different aspects of the same phenomenon, i.e. distribution of the color within the image or region.

In Figure 1 the painting "Map of Israel" from David Reeb is shown. Its XML presentation in MPEG 7 format is given in Figure 2.



**Figure 1. "Map of Israel" from David Reeb**

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<Mpeg7 xmlns="urn:mpeg:mpeg7:schema:2001"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001" xmlns:xml="http://www.w3.org/XML/1998/namespace"
xsi:schemaLocation="urn:mpeg:mpeg7:schema:2001 Mpeg7-2001.xsd">
  <Description xsi:type="ContentEntityType">
    <MultiMediaContent xsi:type="ImageType">
      <Image>
        <MediaLocator><MediaUri>urn:mpeg:image_jpeg:1</MediaUri></MediaLocator>
        <VisualDescriptor xsi:type="ScalableColorType" numOfBitPlanesDiscarded="0"
numOfCoeff="64">
          <Coeff>-8 80 127 -54 4 2 -25 21 18 18 31 32 -7 -3 -46 4 0 11 -3 2 -1 5 1 0 -4 7 -15
0 -3 5 1 -4 2 2 1 12 3 3 5 6 1 1 1 3 3 5 4 5 -4 -3 0 -3 -4 -2 0 -5 0 0 0 -2 -2 -3 -3 -
3</Coeff>
        </VisualDescriptor>
        <VisualDescriptor xsi:type="ColorLayoutType">
          <YDCCoeff>28</YDCCoeff>
          <CbDCCoeff>40</CbDCCoeff>
          <CrDCCoeff>23</CrDCCoeff>
          <YACCoeff5>16 16 14 15 17 </YACCoeff5>
          <CbACCoeff2>16 15 </CbACCoeff2>
          <CrACCoeff2>16 17 </CrACCoeff2>
        </VisualDescriptor>
        <VisualDescriptor xsi:type="ColorStructureType" colorQuant="2"><Values>0 0 0 0 1 7 0
0 8 11 0 3 0 0 2 0 101 157 108 79 0 0 1 1 0 48 108 26 12 144 81 35 255 255 255 255 3 47 104
21 1 52 90 27 2 41 99 50 109 255 248 163 28 95 130 34 0 8 57 45 58 63 26 8</Values>
        </VisualDescriptor>
        <VisualDescriptor
xsi:type="DominantColorType"><SpatialCoherency>1</SpatialCoherency>
<Value><Percentage>2</Percentage><Index>84 100 147</Index><ColorVariance>1 0
1</ColorVariance></Value>
          <Value><Percentage>5</Percentage><Index>130 144 166</Index><ColorVariance>0 0
0</ColorVariance></Value>
          <Value><Percentage>5</Percentage><Index>83 117 117</Index><ColorVariance>0 0
0</ColorVariance></Value>
          <Value><Percentage>6</Percentage><Index>175 195 209</Index><ColorVariance>0 0
0</ColorVariance></Value>
          <Value><Percentage>9</Percentage><Index>66 73 90</Index><ColorVariance>0 0
0</ColorVariance></Value>
          <Value><Percentage>2</Percentage><Index>159 151 152</Index><ColorVariance>1 1
1</ColorVariance></Value>
        </VisualDescriptor>
        <VisualDescriptor xsi:type="EdgeHistogramType"><BinCounts>4 3 5 5 6 1 3 6 2 7 3 3 6
2 7 0 4 3 5 7 4 2 4 6 6 5 2 2 5 7 1 4 5 3 7 5 3 3 6 5 3 3 6 5 6 5 2 4 4 6 5 1 3 5 7 3 2 5 6
6 3 4 5 3 6 3 4 5 4 7 3 3 5 4 7 2 1 3 7 7</BinCounts>
        </VisualDescriptor>
        <VisualDescriptor xsi:type="HomogeneousTextureType">
          <Average>109</Average><StandardDeviation>96</StandardDeviation><Energy>167 169 164
176 190 190 172 181 185 171 186 178 168 187 174 187 187 188 154 175 174 160 183 162 160 167
162 138 170 160</Energy> <EnergyDeviation>167 164 158 174 187 188 160 177 182 170 180 171
153 180 158 175 173 181 142 146 170 134 178 159 143 137 131 121 133 140</EnergyDeviation>
        </VisualDescriptor>
      </Image>
    </MultiMediaContent>
  </Description>
</Mpeg7>

```

**Figure 2. Content of "REEB-MAP OF ISRAEL.MP7.XML"**



Visual attributes can represent global characteristics, concerning whole images, or they can be extracted over the part of the images (specific region or tile of the image). Both approaches have their strengths: global attributes deliver integral temper of the image. Local attributes can capture more detailed information, which characterize the artists' styles or movements' specifics but introduce redundancy for the classifier.

## 4 Vector quantization approach

In [Ivanova et al, 2010] an approach for extracting higher-level visual features for art painting classification based on MPEG-7 descriptors is presented. The complexity of MPEG 7 descriptors as well as the specific similarity measure, which have to be used for them, obstructs their direct use into standard classification algorithms.

The proposed approach consists of the following steps:

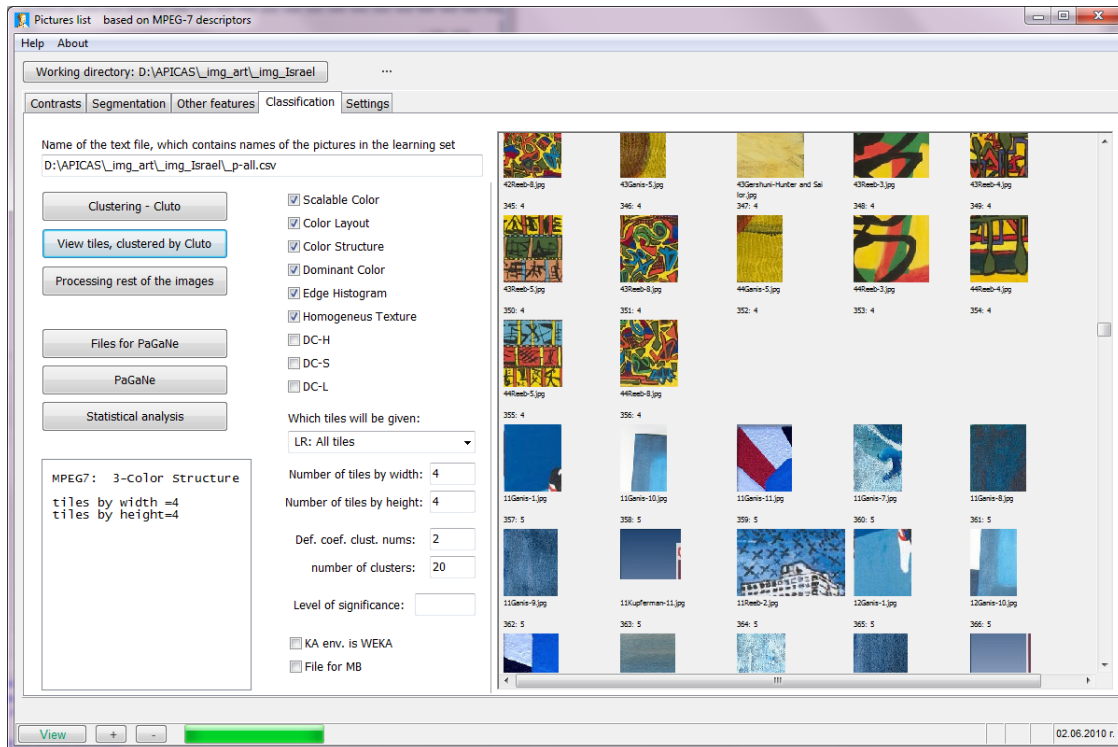
- (1) the images are tiled into non-overlapping rectangles to capture more detailed information;
- (2) the tiles of the images are clustered for each MPEG-7 descriptor;
- (3) vector quantization is used to assign a unique value to each tile, which corresponds to the number of the cluster where the tile belongs to, in order to reduce the dimensionality of the data.

The proposed algorithm is realized as a research tool in the laboratory for art-painting semantic image retrieval "Art Painting Image Color Aesthetic and Semantic" (APICAS). For obtaining the MPEG 7 descriptors [MPEG, 2002] the system refers to Multimedia Content Management System MILOS [Amato et al, 2004]. As a clustering algorithm the program *vcluster*, which is part of the *Cluto* open source software package [Karypis, 2003], is implemented in the system. As a knowledge analysis tool PGN-1 classifier, which is realized in the data mining analysis environment PaGaNe [Mitov et al, 2009] is used.

## 5 Experiments

Here we have used a small dataset that includes 90 paintings of 7 Israel contemporary artists, from Givon Art Gallery ([www.givonartgallery.com](http://www.givonartgallery.com)): Nurit David (23), Maurice Ganis (13), Moshe Gershuni (15), Moshe Kupferman (12), Raffi Lavie (10), David Reeb (12), Pesach Slabosky (5). In brackets the number of images is given. The presence of representative works of contemporary artists is hampered by the existing royalties of the sources, as well as the fact that new artists usually have no structural presence in the digital space yet. Because of this some artists are not well presented.

The experiments were made with 4x4 tiling, using 20 vector quantization values for each MPEG-7 descriptor. In Figure 3 the boundary between tiles, which belong to cluster 4 and 5 for Color structure descriptor, is clearly seen.



**Figure 3. Visualization of tiles grouped by cluster labels for Color Structure Descriptor**

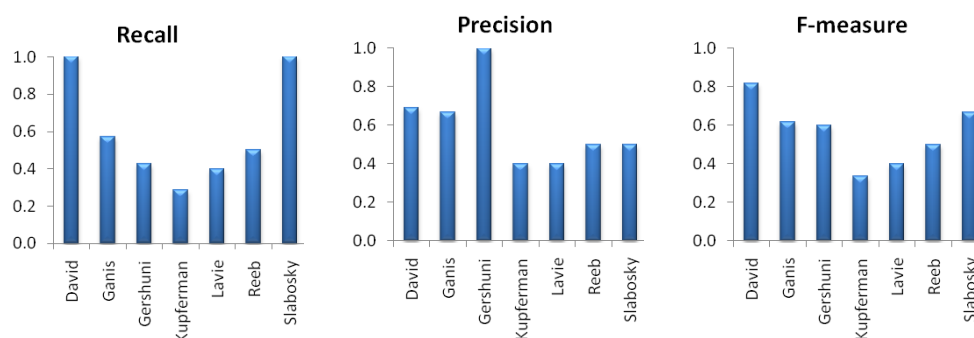
As a result of implementing approach described, each image is represented by a feature vector with 6x4x4 numerical attributes, where 6 is the number of MPEG-7 descriptors and 4x4 is the number of tiles.

The received dataset was split equally into learning and test sets, which are given as input in PGN-1 classifier. The overall classification accuracy was 59.09%. Table 1 gives the confusion matrix of PGN-1 classifier over the dataset. The rows indicate the class where the image actually belongs to. The columns show the class label assigned to the image by the classifier. The numbers of correctly recognized images of the corresponded class are represented on the diagonal.

**Table 1: Confusion matrix of PGN-1 classifier over test dataset**

	Learning set	Test set							Sum by row
		David	Ganis	Gershuni	Kupferman	Lavie	Reeb	Slabosky	
David	12	11							11
Ganis	6	1	4				1	1	7
Gershuni	8	1		3	1		1	1	7
Kupferman	5	1	1		2	2	1		7
Lavie	5	2			1	2			5
Reeb	6		1		1	1	3		6
Slabosky	3							2	2
Sum by column	45	16	6	3	5	5	6	4	45

On the base of this matrix, Recall, Precision and F-measure are calculated (Figure 4).



**Figure 4. Recall, Precision and F-measure for given examining set**

Taking into account the fact that the dataset was not very comprehensive, and the pictures are not sufficient large (most of them were less than 100px width) the results were promising. Analysis of the confusions shows that attributes, extracted from color descriptors such as Color Structure and Dominant Color are very informative. When the style of the author is concentrated more in texture than in color palette the recognition decreases because Homogeneous Texture and Edge Histogram give not so informative attributes.

## 6 Conclusions and future work

In this paper we present tools for indexing and searching in painting digital repositories according to visual characteristics of their content. We have tackled the non-trivial process of image crawling and descriptive feature extraction (we used some of the MPEG-7 features). We study some of the characteristics of the web visual semantics. The

proposed methodology will help us in the transition from Web 2.0 to Web 3.0. Without a breakthrough technology, superior Web 3.0 tools will be more difficult to develop than their counterparts for Web 2.0. This will be part of creation of new tools which will offer society new greater sophistication, complexity, and functionality.

### **Acknowledgement**

This work is partially financed by Bulgarian National Science Fund under the project D 002-308 / 19.12.2008 "Automated Metadata Generating for e-Documents Specifications and Standards".

### **References**

[Amato et al, 2004] Amato, G., Gennaro, C., Rabitti, F., Savino, P.: Milos: a Multimedia Content Management System for Digital Library Applications. Research and Advanced Technology for Digital Libraries: Volume 3232/2004 in Lecture Notes in Computer Science, Springer Berlin/Heidelberg, pp. 14-25 (2004)

[Ivanova and Stanchev, 2009] Ivanova, K., Stanchev, P.: Color Harmonies and Contrasts Search in Art Image Collections. First International Conference on Advances in Multimedia (MMEDIA), 20-25.07.2009, Colmar, France, pp. 180-187 (2009)

[Ivanova et al, 2010] Ivanova Kr., P. Stanchev, E. Velikova, K. Vanhoof, B. Depaire, I. Mitov, Kr. Markov. "Features for Art Painting Classification based on Vector Quantization of MPEG-7 Descriptors". In print.

[Karypis, 2003] Karypis, G.: CLUTO: A Clustering Toolkit Release 2.1.1. University of Minnesota, Department of Computer Science, Minneapolis, MN 55455, Technical Report: #02-017 (2003)

[Maitre et al, 2001] Maitre, H. Schmitt, F. Lahanier, C. 15 years of image processing and the fine arts. In Proceedings of International Conference on Image Processing, Volume 1, pp. 557-561 (2001)

[Mattison, 2004] David Mattison. Looking for Good Art. In Searcher - The Magazine for Database Professionals. Vol. 12, Part I - Number 8, pp. 12-35; Part II - Number 9, pp. 8-19; Part III - Number 10, pp. 21-32 (2004)

[Mitov et al, 2009] Mitov, I., Ivanova, K., Markov, K., Velychko, V., Vanhoof, K., Stanchev, P.: PaGaNe – a Classification Machine Learning System Based on the Multidimensional Numbered Information Spaces. In World Scientific Proceedings Series on Computer Engineering and Information Science, No:2, pp. 279-286 (2009)

[MPEG, 2002] International Standard ISO/IEC 15938-3 Multimedia Content Description Interface – Part 3: Visual,  
[http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=34230](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=34230)

## Authors' Information

Peter Stanchev – Professor, Kettering University, Flint, MI, 48504, USA and Institute of Mathematics and Informatics – BAS; Acad. G.Bontchev Str., bl.8, Sofia-1113, Bulgaria; e-mail: [pstanche@kettering.edu](mailto:pstanche@kettering.edu)

Phillip Ein-Dor – Professor, Head, Information Systems Program, The Academic College of Tel Aviv - Yaffo 4 Antokolsky St., P.O.B. 16131, Tel-Aviv 61161, Israel, e-mail: [philipei@mta.ac.il](mailto:philipei@mta.ac.il)

Krassimira Ivanova – Institute of Mathematics and Informatics, BAS, Acad. G.Bontchev Str., bl.8, Sofia-1113, Bulgaria, e-mail: [kivanova@math.bas.bg](mailto:kivanova@math.bas.bg)

Iliya Mitov – Institute of Mathematics and Informatics, BAS, Acad. G.Bontchev Str., bl.8, Sofia-1113, Bulgaria, e-mail: [mitov@mail.bg](mailto:mitov@mail.bg)