STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **LIII**, Number 1, 2008

# MODELING OF THE IMAGE RECOGNITION AND CLASSIFICATION PROBLEM (IRC)

IOAN ISPAS

ABSTRACT. The problem of the image recognition and classification (IRC) based on the pattern recognition is of a paramount importance in lots of domains. The present paper discusses topics related with the complexity of the algorithms for image recognition and classification. This leads to some precise statements on the computational difficulty of the problem of the image recognition and classification (IRC).

KEY WORDS: modeling, image recognition and classification, algorithm complexity

## 1. Defining the problem of image recognition and classification (IRC)

The automatic classification of the images is of a strategic importance in lots of domains. Its solving is based on the methods and algorithms of automatic pattern/object recognition and image classification [3].

In the following part, we will define the IRC problem:

Given an image data base (data stream) $B = \{I_1, I_2, \ldots, I_n\}$ containing a 'main character', each image incorporating only one object; given a set of descriptions of some known distinct objects $R = \{O_1, O_2, \ldots, O_k\}$; knowing that any human operator is able to recognized easily, by means of rapid visual inspection, the object in the image; the aim of the algorithm is to determine the images that contain these objects, and to classify them in $k+1$ distinct classes: $C_1, C_2, \ldots, C_k$ and $C_{k+1}$. The classes $C_i$, $i = 1, k$ will group all the images containing the objects $O_i$, $i = 1, k$, and the class $C_{k+1}$ will group the images without any of the $R$ objects.

The diagram of the IRC problem is the following:

Received by the editors: April 2, 2008.

2000 *Mathematics Subject Classification.* 93A30, 68T10.

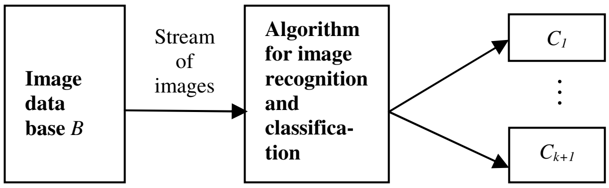
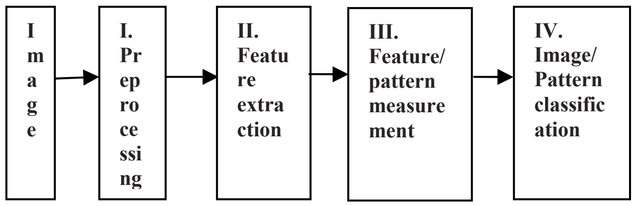



Figure 1. offers a small example consisting of eight images with 'characters' [13] that can be classified into seven image classes: 1. the class of the images containing horses; 2. the class of the images containing cheetahs; 3. the class of the images containing elephants; 4. the class of the images containing airplanes; 5. the class of the images containing bears; 6. the class of the images containing eagles; 7. the class of the images "neutral", without recognized object. In this case, the component elements of the set of the recognized 'objects' are: 1. horses; 2. cheetahs; 3. elephants; 4. airplanes; 5. bears; 6. eagles.

FIGURE 1. Examples of images to be classified

Having studied the specialized literature [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11] we can state that the image recognition algorithms describe a four-stepped process. Each step is essential and inevitable. The diagram of the image recognition process is the following:




**I. Preprocessing of the image.** This means the application of some DIP (Digital Image Processing) algorithms specialized in enhancing image quality [1], [3], [5], [10].

**II. Feature extraction.** This is the key step, the one that measures the performances and the quality of the recognition software. The discovery of the most proper features and characteristics of the recognition object is the key of the success [12], [14], [15], [16], [17], [18], [19], [20]. The $FeatureExtraction$ algorithm, the implementation of this essential step, output a feature vector description of the recognized object $(v_1, v_2, \ldots, v_n)$, not necessarily numerical.

**III. Feature/pattern measurement.** This step is well theoretically founded; there is a developed mathematical theory (The measure theory) which can help us select the proper and efficient n-dimensional metrics. The final result of this step is usually a one- or multi- dimensional value (a vector) percepted as the 'distance' of the feature vector towards the borders of the class [1], [12], [15], [18], [19].

**IV. Image/Pattern classification.** This is the final step which combines the results of the prior measurements. The pattern/object - described by the feature vector - belongs to a class of images, according to certain appartenance mathematical criteria. The final result of the classification step must be the index $i$ of the image class $C_i$ .

## 2. The modeling of the problem of the image recognition and classification

The primary modeling of the problem of the image recognition and classification is an extremely difficult subject. In order to reduce its degree of difficulty, a gradual approach is indicated to be used in a step-by-step manner.

### 2.1. **The simplified version of the problem of the image recognition and classification (sIRC).**

If we consider the objects-'characters' from the images as marks/signatures, that were previously inserted in the images. Consequently, every object became obviously a "main character", in front of the scene. Then the recognition of images leads us to the following simplified version of the IRC problem defined at the beginning:

Given an image data base (or an image stream) $B = \{I_1, I_2, \ldots, I_n\}$ containing a 'main character' that marks them; given a set of descriptions of some known distinct objects $R = \{O_1, O_2, \ldots, O_k\}$; the following algorithm determines the images from B that contain these objects and classifies them in $k+1$ distinct classes: $C_1, C_2, \ldots, C_k$ and $C_{k+1}$. The algorithm has an image $I$ as input and is calling the sub-algorithm *Recognition*; this algorithm decided if object $O_k$ is contained in image $I$.

Algorithm **sIRC**(image I);
For $i = 1, k$ do
    If **Recognition** $(O_i, I)$ return $(i)$;
Return $(k + 1)$;

Our belief, just like its title shows, is that the simplified version **sIRC** problem is less difficult than the initial one. Unfortunately, we cannot prove rigorously this statement although the multitude of facts strongly confirms it. It is obvious that its complexity relies on the complexity of the Recognition sub-algorithm. The total complexity of the algorithm is in the worst of cases:

$$\text{WorstCase}(\textbf{Classification\_sIRC}) = k \times \mathcal{O}(\textbf{Recognition}),$$

where $k$ is the dimension of the set of objects $R$. $\mathcal{O}(\textbf{Recognition})$ is the classical notation for the complexity class of the Recognition algorithm.

The **Recognition** algorithm is the clue of the **sIRC** problem. Its input is I image and the description of the recognition pattern/object O. For every object $O_i$ it works like a validation function with the output true or false. Considering that any human operator is able to easily recognize the presence of the object O in the image I by visual inspection, based on a primary process of the mathematical modeling and formalization of the sIRC problem, we can design the following modeling:

### 2.2. **Mathematical modeling of the sIRC problem.**

**Definition 1.** A **searching space** is a set of data $S$ which has to be exhaustively covered in order to find the target data $x$ among $S$ data.

Giving $n$ the cardinal of the set $S$; considering that the exhaustive covering condition is needed, then the number of required steps (comparisons) in order to find out $x$ is in the worst case $n$.

**Definition 2.** A **pattern** $P$ of an ($b_i$-dimensional) object is the set of the contour points (laying on the external edges of the shape of the object) which delimits the space occupied by it.

The pattern $P$ of an object is that what makes it distinguished from the environment and confers its identity.

**Definition 3.** An **informational content** (colorist) $C$ of a certain object is a set of points belonging to the object, grouped together according to an association (relational) criteria.

For instance, the set of the 'interior' points of the object, the set of the points of the same color, etc. The information content (colorist) C is the visible, descriptive expression of the object.

**Proposition 1.** Therefore we can state that every object is uniquely defined by its pattern $P$ and its informational content $C$. As the pattern $P$ and its informational content C are described by numerical vectors, the pair $(P, C)$ uniquely defines every object.

### 2.3. **Introducing the auxiliary problem IRC(R).**

The terms of the auxiliary problem IRC(R) are the following: this problem is particular case of the sIRC problem in which the set of objects to be recognized consists in a single type of objects $R$, as an image of the solid rectangle.

Any rectangle $R$ is defined by the pair of corners $A(x_A, y_A)$ and $B(x_B, y_B)$, and by its colour $C$. Thus $R = R(A, B, C)$.

The problem requests to determine the subset $Q$ of the image set containing one rectangle $R$.

The easiest method (considering the effort in designing the recognition algorithms) is the scanning of all the images, i.e. for every image, to check up every possible matching position of the rectangle $R$. The effort of recognition and classification of the images, which is directly proportional to the image resolution and indirectly proportional to the dimension of the rectangle $R$, will thus be huge. Redesigning this brute force approach method implies of new, more efficient methods to match the rectangle $R$, other than exhaustive scanning of the image. Since an image $I$ certainly contains a rectangle $R$, one can question if the image $I$, having $M \times N \times c$ resolution, could be seen as a searching space for rectangle $R(A, B, C)$, where $c$ is the color resolution of the image $I$.

Given the set of all valid coordinations $S(A, B)$ for the rectangle $R$ in the image $I$. Every recognition algorithm $\mathcal{A}$ of $R$ in $I$, seen from the **Turing-Church Thesis** perspective [21], can be simulated by a Universal Turing Machine $U$ is possible. The algorithmic complexity of the Machine $U$ is identical with the recognition algorithm $\mathcal{A}$. Moreover the Universal Turing Machine $U$ is designed to display every pair of possible coordinates $(A, B)$ as algorithm $\mathcal{A}$ overpasses them at its run-time, following its specific steps.

Therefore, it is obvious that the set of all valid coordinations $S(A, B)$ becomes a searching space for the algorithm $\mathcal{A}$. None of the pairs $(A, B)$ can avoid being checked. In conclusion, the difficulty of the $IRC(R)$ problem is the same with the difficulty of finding a value $x$ in the searching space $S$ by exhaustively scanning. In this case, $x$ represents the coordinate pair $(A, B)$ while $S$ represents the set of all valid coordinations the rectangle can have in the image.

2.4. **Reduction of the sIRC problem to IRC(R) problem.** According to Proposition 1 every recognized object $O$ is defined by the pair of vectors $(P, C)$, pattern and informational content. Since $P$ and $C$ are vectors and not singular values, the pair $(P, C)$ denotes a rectangle in a multidimensional space. The recognition of the object $O$ in the set of the images can be consequently regarded as the matching of the rectangle $(P, C)$ in the corresponding searching space, resulted from the union of all the pairs of valid coordinates $(P, C)$ of the object $O$ in the images. Thus, the sIRC problem is reduced to the IRC(R) problem.

## 3. Conclusions resulting from the mathematical modeling of the sIRC problem

**Conclusion 1.** Generally speaking, for each Recognition$(O, I)$ algorithm the image $I$ becomes a searching space for the object to be recognized $O(P, C)$.

**Conclusion 2.** The complexity of the algorithm Recognition$(O, I)$ is direct proportional with the dimension of the image $I$ and with the dimension of the codes $P$ and $C$ :

$$\mathcal{O}(\text{Recognition}(O, I)) = \mathcal{O}(Dim(I) \times Dim(P) \times Dim(C))$$

## 4. Final results and conclusions

Resuming the statements before, we can reach the following final results.

**Result 1.** The difficulty of the simplified version of the problem sIRC is a consequence of the complexity of the algorithm Recognition$(O, I)$.

**Result 2.** The essential step in the recognition process of the object $O$ is the feature extraction of the patterns and of the information content $(P, C)$ from the image $I$. Notice. This feature extraction step is inevitable because the image $I$ is formed by a matrix of pixels, but the descriptors $(P, C)$ are a pair of codes describing the shape and the information content of the object $O$.

**Result 3.** The complexity of the algorithm Recognition$(O, I)$ is directly proportional with the dimension of the features $(P, C)$ extracted by the sub-algorithm FeatureExtraction$(O)$, the main component of the second step of the recognition process, pointed out earlier in the diagram of the image recognition process.

**Result 4.** The complexity of the algorithm Recognition$(O, I)$ is given by the formula:

$$\mathcal{O}(\text{Recognition}(O, I)) = Dim(I) \times \mathcal{O}(FeatureExtraction(O))$$

where $\mathcal{O}(FeatureExtraction(O))$ is the complexity class of the $FeatureExtraction$ algorithm.

**Important notice**. The extraction of the color and of the pattern features from the image may imply a very consistent number of operations $m$ (i.e. associations and relations) over the pixels within the interest zones. The complexity of the extraction algorithm of the object $O$ becomes:

$$\mathcal{O}(\text{FeatureExtraction}\,(O))) = m \times Dim(\text{ExtractionZone})$$

Note that the determination/discrimination of the interest zones (which could contain the object) is the most important but also the most difficult step in the entire feature extraction process. This can lead to a situation wherein the recognition of an object having dimension $200 \times 200$ pixels, within an image having a resolution of $800 \times 600$ pixels and 256 colors, could require a number of operations directly proportional with the huge value $800 \times 600 \times 256 \times 200 \times 200$, greater than $10^{12}$.

The final conclusion about the difficulty of the image recognition and classification problem is that a proper solution of the problem and of its simplified version sIRC depends in the most direct way on the design of an efficient pattern/information content extraction sub-algorithm.

## References


[1] Gonzalez R., Woods R. - Digital Image Processing, Prentice Hall, 2002, 2nd Edit.

[2] Jain A., Duin R., Mao J. - Statistical Pattern Recognition: A Review, IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, pp. 720-736, January 2000.

[3] Russ, John C., The image processing handbook, 5th ed., CRC Press, 2006.
[4] Richard O. Duda, Peter E. Hart, David G. Stork, Pattern Classification, 2nd ed., Wiley Interscience, 2000.
[5] Sankar K. Pal, Amita Pal (eds), Pattern Recognition. From Classical to Modern Approaches, World Scientific Publishing Company, 2002.
[6] Mitra Basu and Tin Kam Ho (eds), Data Complexity in Pattern Recognition, Advanced Information and Knowledge Processing, Springer-Verlag, 2006.
[7] Bahram Javidi (ed), Image Recognition and Classification. Algorithms, Systems, and Applications, CRC Press, 2002.
[8] Bernd Jahne, Digital Image Processing, 5th ed., Springer, 2002.
[9] Yali Amit, 2D Object Detection and Recognition. Models, Algorithms and Networks, The MIT Press, 2002.
[10] Sing-Tze Bow, Patern Recognition and Image Preprocessing, 2nd ed., Marcel Dekker Ltd., 2002.
[11] Ioan Ispas, The image recognition and classification, a four-steped modeling, Proc. 2nd International Conf. on European Integration - Between Tradition and Modernity, Petru Maior University, Tîrgu Mureş, Sept 20-21, pp. 724-730, 2007.
[12] Kian-Lee Tan, Beng Chin Ooi, Chia Yeow Yee - An Evaluation of Color-Spatial Retrieval Techniques for Large Image Databases, Multimedia Tools and Applications, 14, 55-78, 2001, Kluwer Academic Publishers.
[13] elib.cs.berkeley.edu/photos/classify/
[14] Oge Marques, Borko Furht - Muse: A Content-Based Image Search and Retrieval System Using Relevance Feedback, Multimedia Tools and Aplications, 17, 21-50, 2002, Kluwer Academic Publishers.
[15] Y. Alp Aslandogan, Clement T. Yu, Ravishankar Mysore, Bo Liu - Robust content-based image indexing using contextual clues and automatic pseudofeedback, Multimedia Systems 9: 548-560 Springer-Verlag 2004.
[16] Kian-Lee Tan, Beng Chin Ooi, Chia Yeow Yee - An Evaluation of Color-Spatial Retrieval Techniques for Large Image Databases, Multimedia Tools and Applications, 14, 55-78, 2001, Kluwer Academic Publishers.
[17] A. Srivastava, A.B. Lee, E.P. Simoncelli, S.-C. Zhu - On Advances in Statistical Modeling of Natural Images, Journal of Mathematical Imaging and Vision 18: 17-33, 2003 Kluwer Academic Publishers.
[18] Jörg Dahmen, Daniel Keysers, Hermann Ney and Mark Oliver Güld - Statistical Image Object Recognition using Mixture Densities, Journal of Mathematical Imaging and Vision 14: 285-296, 2001, Kluwer Academic Publishers.
[19] Martin Heczko, Alexander Hinneburg, Daniel Keim, Markuswawryniuk - Multiresolution similarity search in image databases, Digital Object Identifier (DOI) 10.1007/s00530-004-0135-6, Multimedia Systems 10: 28-40, Springer-Verlag 2004.
[20] Wei-Ying Ma, B. S. Manjunath - NeTra: A toolbox for navigating large image databases, Multimedia Systems 7: 184-198 (1999) Multimedia Systems, Springer-Verlag, 1999.
[21] Stanford Encyclopedia of Philosophy, The Church-Turing Thesis, first published Jan 8, 1997; substantive revision Aug 19, 2002. http://plato.stanford.edu/entries/church-turing/

N. Iorga 1, Tîrgu-Mureş
*E-mail address*: `ispas@science.upm.ro`