Contributions to nonlinear and adaptive processing of JPEG compressed images

Scientific coordinator:
Professor Aurel VLAICU, PhD

2009
Contents of the thesis:

1. Introduction
2. Using the information available in DCT coefficients for image analysis
   2.1. Discrete cosine transform
   2.2. A proposed algorithm for detecting/localization and hiding patient identity information in ultrasound images
   2.3. The relations between a DCT coefficient block and its sub-blocks
   2.4. Support vector machines in the visual recognition of objects in the compressed domain
      2.4.1. Support vector machines
      2.4.2. Dot product of two vectors in the original feature space and in a discrete cosine transform – induced feature space
      2.4.3. Implementation of the training and classification phases in the compressed domain
         2.4.3.1. A standard face recognition application
         2.4.3.2. An application for object recognition and localization in the compressed domain
3. Implementation of nonlinear operations for image processing in the compressed domain
   3.1. Linear and nonlinear operations in the compressed domain
   3.2. A general adaptive solution to the problem of brightness thresholding
   3.3. Image binarization in the compressed domain
   3.4. The fuzzy intensification operator for contrast enhancement in the compressed domain
   3.5. Compressed domain image enhancement algorithm based on a Takagi-Sugeno fuzzy inference system
4. Image segmentation in the compressed domain
   4.1. A Bayes classifier based approach for image segmentation in the JPEG compressed domain
   4.2. Applications to morphometrical analysis of liver biopsies
5. Conclusions; future research work
Introduction

The topic addressed in the doctoral research refers to the compressed domain processing of JPEG compressed scenes / MPEG compressed videos.

With the increasing sizes of high resolution images, their storage and processing directly in the compressed domain has significantly gained importance. The bitmap format for image storage is seldom used in day-to-day applications; even medical images are stored in JPEG format. This is because of the advantages offered by the JPEG format: little storage space needed and better performances in information transmission (via internet). For images stored in JPEG format, it is recommended to process them directly in the compressed domain, in order to reduce the time needed to process data. First, this time economy is related to the fact that it is no longer necessary to decompress the image, to process it at pixel level and to recompress it back; and, second, the image processing in the compressed domain means that there are fewer data to process.

If in the case of image processing at pixel level the new algorithms try generally to provide solutions with better results in quality than the existing ones, in the case of processing in the compressed domain the main aim is to achieve the same quality as in pixel level algorithms but at the same time to improve the numerical efficiency of processing, compared to the case of JPEG decompression based processing. As such, all the methods in this field of research are compared with pixel level methods and tend to achieve similar performance, and not necessarily an improvement in the quality of results.

The personal contributions in this thesis are the theoretical development of new algorithms and deployment strategies for existing algorithms, new systems based on existing mathematical principles and the application of these algorithms and systems to certain problems of image analysis.

The contributions presented in this thesis follow the main trends in the state of the art of this field of research, are theoretically explained and experimentally verified. They are also validated by the publishing of nine scientific papers in international conferences, journals and one book chapter in the field. Some contributions are included in the reports of two nationally funded research grants.

Thesis structure

The thesis is structured in five chapters and a references chapter.

Chapter one is an introduction. It includes a short presentation of the research trends in processing JPEG images in the compressed domain, a description of the principle behind image processing in the compressed domain, the advantages associated with this type of processing and a short presentation of the important works published (state of the art). It also describes the basic principles of the JPEG standard and the JPEG file format.

Chapters 2, 3 and 4 consist of the detailed description of the personal contributions in the field of nonlinear and adaptive processing of JPEG compressed images by developing new solutions for:

- analysing/processing JPEG images in compressed format leading to applications that in particular exploit the information found in DCT coefficients of the image blocks and are extremely useful for various practical situations (chapter 2)
- the compressed domain implementation of some highly nonlinear grey scale transformations, such as brightness thresholding, and for numerically efficient implementation in the compressed domain of some contrast enhancement algorithms with
Contributions to nonlinear and adaptive processing of JPEG compressed images

- the implementation in the JPEG compressed domain of some advanced algorithms for image analysis and segmentation based on supervised classifiers (chapter 2, chapter 4).

Each of these three chapters begins with a section presenting the general overview of the image processing class, the practical applications of the presented contributions and the theoretical background for these contributions. Each chapter contains at least a subchapter dedicated to the detailed and complete presentation of one of the contributions (newly proposed algorithm/system) starting with the basic ideas on which the new solution was built, continuing with the mathematical support and finishing with the experimental verification of the solution on standard sets of test images and the evaluation of the performance compared with other similar solutions in literature.

Chapter 2 presents the personal contributions regarding direct analysis of the information available in the DCT coefficients of the elementary 8x8 pixels blocks which form the basic units of the JPEG image. The chapter begins with a presentation of the general framework of the discrete cosine transform and of the information carried by the DCT coefficients (subchapter 2.1), as well as an approach for obtaining the relations between a DCT coefficients block and its sub-blocks (chapter 2.3).

Subchapter 2.2 outlines a method proposed for analysis and segmentation of medical ultrasound images, using the information contained in the RLE (Run Length Encoded) vectors, as a solution for processing/analysing the JPEG compressed images. These types of applications, which use the information contained in the compressed format immediately after the Huffman decoding, are extremely useful for certain practical situations. This type of approach could open up a new field of use for compressed domain algorithms, providing solutions that directly exploit the compressed representation and are easier to implement in this form rather than using the spatial pixel representation. In this context, an algorithm for detecting/localization and hiding patient identity information in ultrasound images directly in the JPEG compressed domain [Popa08c] is proposed, using the energy contained in DCT coefficients.

Subchapter 2.4 presents the contributions in using support vector machines in the visual recognition of objects in the compressed domain, by directly using the information contained in quantized DCT coefficients. The motivation behind the choice of support vector machines (subchapter 2.4.1) to implement some detection/recognition systems in the JPEG compressed domain is that through their mathematical principle, they are characterized by the ability to learn from a relatively small number of examples (meaning that they have a very good generalization capacity), and by the ability to separate highly correlated data by projecting them in a higher dimensional space in which the data becomes linearly separable.

The main mathematical operation between the feature vectors in the classification phase of a support vector machine is the computation of the dot product of two vectors. In subchapter 2.4.2 it was proven mathematically that the dot products is the same in both the original feature space and in a unitary transform induced feature space, such as the one induced by the discrete cosine transform which is the base for JPEG compression. As such, the implementation of support vector machines in the JPEG compressed space, with quantized and zigzag ordered DCT coefficients vectors is feasible and it can be expected to lead to similar performances like the application of support vector machines in the original data space (the differences in the classification performances are due to the quantization, but in some cases this will be an advantage for the classification). The proposed algorithm was verified practically by implementing the training and classification phases in the compressed domain for a standard face recognition application (subchapter 2.4.3.1). The analysis was achieved by taking into account a set of quantized greyscale JPEG facial images of a minimum possible dimension (10x8 pixels) so that they retained enough details for a successful recognition.

Subchapter 2.4.3.2 describes the theoretical formulation and practical implementation of
a support vector machine classifier with binary decision function for object localization in a JPEG image, considered decomposed in a set of partially overlapping rectangular windows of any dimension (obtained by scanning the image with a rectangular window of an a-priori chosen size, with a randomly chosen scanning step - measured in pixels). To obtain the quantized DCT coefficients feature vectors for any given size window (corresponding to the object that is to be recognized), starting from the 8x8 DCT coefficient blocks used in JPEG encoding, it is necessary to use the relations between a DCT coefficient block and its sub-blocks – presented in detail in subchapter 2.3. The mathematical background of support vector machines and the reasons for applying them for object recognition in JPEG compressed domain data are presented in the research report [Popa08d]. The validation of the proposed implementation, by using a binary support vector machine classifier, both linear and nonlinear, was achieved by an application for eye detection and localization in JPEG compressed color face images. The performances achieved, compared to the application based on pixel data, support the theoretical reasoning. In the case of using quantized DCT coefficients, the performances are slightly superior to the ones based on the spatial representation of the image.

The main advantage of the developed system is given by the possibility of using the available information straight from the JPEG compressed images (the DCT coefficients), without the need for decompressing the digital image to pixel level in order to apply the support vector machine classifier. Another advantage is that, by using a feature space with a very good energy compacting property, a large number of values in the feature vectors will be zero, thus significantly reducing the number of multiplications and summations involved in the computation of the decision function of the support vector machine. The reduction of the computational complexity is further emphasized by using the quantized DCT coefficients. By reducing the feature vectors length by retaining only the most significant values in the top left corner of the quantized DCT coefficients block, the rate of correct recognition drops insignificantly, but the feature vectors are significantly shorter (by keeping only half of the columns and lines, and thus only a quarter of the values), thus increasing again the computational speed.

Chapter 3 presents the contributions in the adaptive implementation of nonlinear operations for image processing in the compressed domain. Chronologically, the first types of image processing algorithms implemented in compressed space were simple arithmetical pixel operations which involve only operations like adding, subtracting, multiplying and dividing with a constant or pixel adding. In this case the operation made in the compressed space is identical to the one in the spatial domain due to the linearity of the discrete cosine transform (subchapter 3.1). At an immediately superior complexity level we find the algorithms that require a pointwise multiplication of the luminance between two images, a more complex operation in the case of the implementation in the compressed domain, because it requires the computation of a convolution product; a particular case of such an operation is computing the power of the luminance in the image pointwise, for a certain value of the exponent (i.e., square, cube).

Actually, in general, the more nonlinearity the spatial operation at the pixel level involved in the algorithm has, the more difficult it is to implement it in the compressed domain. Paradoxically, the simplest algorithms of image processing at the pixel level, such as thresholding (image binarization), the greyscale clipping contrast enhancement (piecewise linear), become the most difficult to transpose in the compressed domain because they involve severe nonlinearities. This is the reason for which, at present solution are sought for efficiently implementing these types of operations.

Given this situation we formulated a general adaptive solution for the problem of luminance thresholding in a digital image with a minimum decompression of the JPEG image (subchapter 3.2). This solution is adaptive for the content and distribution of luminances in the 8x8 pixel block that represents the base unit in JPEG compressed images (distribution reflected
Subchapter 3.3 describes a proposed application for image binarization in the compressed domain, which is a very simple operation at pixel level but very difficult to transpose in the compressed domain [Popa06]. The numerical efficiency of the proposed algorithm, in the context of an insignificant (negligible) error in respect to the classical processing based on complete decompression of the JPEG image and thresholding at pixel level, is demonstrated by experiments on a significant number of standard test images.

The proposed solution was further applied in the threshold comparison steps involved by two fuzzy algorithms dedicated to contrast enhancement directly in the compressed domain:

1. A contrast enhancement algorithm of an image based on unary fuzzy sets operations – using the fuzzy intensification operator INT (subchapter 3.4) [Popa07, Florea09a, Florea09b]. This algorithm involves a threshold comparison and the computation of the square of the luminance pointwise.

2. An image enhancement algorithm based on a Takagi-Sugeno fuzzy inference system – system based on fuzzy rules (subchapter 3.5) [Popa08a, Popa08b, Florea09b]. This contrast enhancement algorithm, unlike the previous one which uses the fuzzy INT operator, involves three threshold comparisons for the selection of the correct linear function to be applied for the intensity values in a block, thus increasing the degree of nonlinearity.

These algorithms are widely used in the image processing literature because of their performances and easy adaptive parameter choice based on the greyscale histogram of the image, but are characterized by a high degree of nonlinearity, which explains the lack of such fuzzy algorithm formulations in the compressed domain. For the selection of the parameters of the membership functions directly in the compressed domain, the use of the histogram of the DC coefficients of all DCT coefficient blocks in the image as an approximation of the statistic of greyscale levels in the image is proposed (taking into account the fact that in the compressed domain we cannot access the greyscale image histogram).

The most difficult to transpose nonlinearity is the threshold comparison; in this way, the use of the proposed and verified solution through the previously mentioned contribution allowed for a relatively easy formulation of fuzzy algorithms in the compressed domain. The luminance processing errors due to the threshold comparison of only the average luminance in every block in the JPEG image are less visible, given that the processing result is a greyscale image (unlike the more restrictive case of binarization). This allows for a higher numerical efficiency in the above mentioned contrast enhancement cases.

Each of these subchapters that contain contributions based on the implementation of threshold comparison, includes a subsection that presents the basic principles of the classic algorithm (at the pixel level), having the aim of clarifying the notions needed for understanding the reformulation of the algorithm directly in the compressed domain.

Chapter 4 presents the last contribution in the thesis: the implementation of an advanced analysis and segmentation algorithm of color digital JPEG images based on the Bayes classifier with Gaussian probability model directly in the compressed domain. It can be estimated that this field will be interesting for the scientific community in the future, given the implicit advantage offered by the information existent in the compressed domain, with no preliminary extraction of features – texture information offered by the DCT coefficients and the decorrelation of the color information by using the YUV space.

The theoretical formulation and practical implementation of the Bayes classifier with Gaussian probability model, dedicated to image segmentation based on color and texture attributes, is described in the subchapters 4.2 and 4.3 of the thesis. In the proposed approach, both the training data and the data to be classified are the quantized DCT blocks represented as
vectors in the YUV color space. Because discrete cosine transform is unitary, the equivalence of Bayesian classification with Gaussian probabilities models in the domain of DCT coefficients and at a pixel level can be demonstrated. The quantization of DCT coefficients (implicit in the JPEG standard) and the use of the YUV color space (in which the color information is decorrelated from the brightness information, unlike in the RGB primary color space) provides advantages in the classification, by enhancing the relevant information for discrimination of different image regions and discarding the irrelevant information which could lead to false segmentation results. Therefore the correct classification rate using the proposed algorithm (as the experiments show) in a proposed medical imaging application is superior to the classification in the spatial domain.

The proposed Bayesian segmentation solution is applied to the fast and accurate identification and quantification of the hepatic fibrosis from liver biopsies, using Gaussian models for the categories of tissues [Popa08d, Gordan08]. The approach is based on the pixel level color information, but also on the local texture information in 8×8 pixel neighborhoods, making thus use of the already available information in the microscopic images at hand by considering their representation in the JPEG format. Thus, the image storage format itself provides the information needed for an accurate identification/segmentation of the hepatic biopsies into tissue vs. microscopic slide and further, of the already identified tissue areas into healthy tissue vs. hepatic fibrosis. The experimental results of the proposed method prove not only a good accuracy but also a fast classification, due to the benefits of the transform domain data, which makes it sufficiently to perform a block classification rather than a pixel level classification. Whereas complicating to some extent the training phase, the classification (test) phase is significantly faster. Also the accuracy provided by the implicit use of texture information is superior to the one based by pixel-level color information alone.

The last chapter of the thesis, chapter 5, formulates the conclusions derived from the research work performed during the doctoral studies (the results of which are presented in chapters 2, 3, 4), examined the contributions made, their relevance to the state of the art, their performances and the improvements to be made to the proposed algorithms and systems. The last part of the chapter is dedicated to the future research work regarding the further development of the presented algorithms.

**Main personal contributions**

I theoretically developed, implemented and validated through software implementation and several experiments (on standard data sets):

1. An application for identifying personal data of patients in JPEG compressed ultrasound images [Popa08c]
2. The use of support vector machines in visual recognition of objects in the compressed domain [Popa08d]
3. The development of a new method of implementing luminance threshold comparison type nonlinearities, with the following applications:
   3.1. An image binarization algorithm of images in the compressed domain [Popa06]
   3.2. A new method of image enhancement in the compressed domain using the Fuzzy INT operator [Popa07, Florea09a, Florea09b]
   3.3. The formulation in the compressed domain of an image enhancement algorithm based on a Takagi-Sugeno fuzzy inference system [Popa08b, Florea09b]
4. A new algorithm for image segmentation using the Bayes classifier with Gaussian probability model directly in the compressed domain [Popa08d, Gordan08]
Author’s Publications Referenced in this Thesis


