Image Processing Filters Based on Image Analogies

Jelena Z. TASIĆ                                    Milan TUBA
Faculty of Mathematics                    Faculty of Computer Science
University of Belgrade                Megatrend University Belgrade
Studentski trg 16, Belgrade          Bulevar umetnosti 29, N. Belgrade
SERBIA                          SERBIA
jelenatasicsd@yahoo.com                          tuba@ieee.org

Abstract: - This paper presents a method to develop a single image processing filter that can restore images subject to multiple distortions. Often medical images originated by alternative sources like ultra-sound or x-rays or images from cameras on remote locations are subject to multiple distortions like damaged optics, undesirable movements or natural noise. Information about exact sources of distortions and their parameters are usually not available. Proposed method is based on image analogies and it requires one correct image before distortions and one image of the same object under distortions. The feedforward architecture of neural networks and the resilient propagation learning algorithm are used to generate a single filter that successfully reverses all distortions and can be applied to all images from the same distorted source. On different test examples our method achieved very good results so it can be considered a promising tool for such corrections.

Key-words: Image processing, Image analogies, Neural networks, Filters

1 Introduction
Digital image processing is among most important research topic because of its many applications. Digital images are nowadays used almost everywhere and in most cases some image processing is required. In many applications it is of crucial importance since without it images would be useless. Examples are many medical images originated by alternative sources like ultra-sound or x-rays or images from cameras on remote locations where it is impossible to correct problems like damaged optics or undesirable movement on the spot. If we know the sources and the nature of the distortions we can design mathematical tools, usually filters, to counter mentioned distortions. However, if there are many such distortions and we do not have precise information about their nature and parameters, the problem becomes more difficult.

In this paper we propose a method of designing image processing filters based on image analogies. If we have one „good” image, before distortions and the same image with distortions, we develop a filter using neural networks that transforms the distorted image to the original without distortions. A single filter reverses all the distortions without separating them or even trying to determine them. Such a filter can later be used to process all images from the same source.

2 Related work
General idea of this paper – learning filters on the basis of a given sample – is based on Hertzmann’s article [1] where the term image analogies is introduced: „Given a pair of images $A$ and $A'$ (the unfiltered and filtered source images, respectively), along with some additional unfiltered target image $B$, synthesize a new filtered target image $B'$ such that:

$$A : A' = B : B'$$ (1)

The similarity metric is based on approximation of Markov random field model and uses pixel values for learning of a chosen filter. The statistics of the selected pixel neighborhoods is used for calculating the relationship between the original and resulting pair of images. This algorithm uses a multi-layered representation of the images $A$, $A'$ and $B$ (Gaussian pyramid), together with a feature vector as well as the approximate nearest neighbor method for speeding up the mapping process. Hertzmann uses the same framework for curve analogies [2], learning of curve style on the basis of a given example.

As opposed to Hertzmann’s algorithm [1] which enables learning of different kinds of filters, other algorithms specialize in narrower spectrum of filters but hoping to achieve better results. Freeman proposed an algorithm [3] which offers a way of guessing a range of high frequencies that are missing in the image. Algorithm takes training images as references with the aim of learning image sharpening. For the training data the algorithm uses
a set of images of low resolution and the corresponding images of high resolution. The algorithm [4] solves the problem of improving the quality of personal photographs by using the favorite set of personal photographs as a training set. The algorithm [5] is using Hertzmann’s idea of image analogies where a pair of images is used for training. This algorithm uses semi-supervised learning technique, Markov’s random model in order to secure global coexistence, and image quilting technique in order to secure a local coexistence. In the article [6] a method that improves a visual quality of satellite images by using the examples of super-resolution techniques is described.

Greyscale coloring of the images is a difficult problem. Levin [7] proposes a method based on a simple assumption: neighboring pixels in the space that have a similar intensity should also be of similar colors. This algorithm does not require the precise segmentation of the image. It is necessary that the user marks the color of each region of the image that is to be colored. As opposed to this approach Irony’s [8] algorithm finds several points on the image which the algorithm had colored correctly. Then Levin’s approach is applied, as though the points are those that the user marked.

In medical industry a substantial amount of data is being collected, while only a small number can be correctly analyzed. There is no reliable way of quickly segmenting a large amount of data. In the article [9] image analogies are used for resolving the problem of image segmentation. Medical images reviewed by the experts are used as training data.

In the article [10] a method is proposed for the automatic improvement of brightness and contrast by using supervised learning.

3 Implementation and Methodology of the Proposed Solution

Analogy is the basic reasoning process. People often use analogies without usually even being conscious about it. They use them for predicting or resolving different kinds of problems. Generalization based on a set of familiar examples is the central problem of machine learning. The basic premise of this paper comes from the essential idea of analogy presented by Hertzmann [1]. The idea of this work is to use neural networks for learning analogies.

3.1 Learning Technique and Training Data

As a program input two images are taken: the original image (model) and the filtered image (master). Neural networks learn model-master mapping, and after a given number of generations, the achieved result is displayed in the apprentice window.

For choosing a training data set two options are enabled: a choice of a number of a random set of pixels from the input images (model and master) or the manual selection of image surfaces to be used as training data. The colors on and around any given pixel p of the image model correspond to the colors on and around that same pixel p of the image master. It is necessary to select the training data set carefully since it does not contain specific information, the desired result may not be achieved.

The learning procedure takes the achieved and the desired results and compares them, after which the network connection weights are corrected. The same set of data is processed several times during the network training until the connection weights are improved. The procedure is repeated until the desired similarity - fitness condition is satisfied when the program remembers the determined neural network.

3.2 Architecture

It is necessary to select the architecture, activation function and the algorithm for training the neural networks.

Feedforward neural networks have been chosen, which consist of the input, the hidden and the output layer of neurons, where the upper layers do not supply feedback to the lower ones. The network is structured in such a way that each neuron of the hidden layer receives the signals from all the neurons from the previous layer.

In addition to the pixels chosen as training data, the size of the retina is also given as a parameter which influences the number of neurons in the input and the hidden layers. If this parameter has value 3, the size of the retina is 3x3, the number of neurons in the input layer for a color image is 3*9, because a color image has 3 bands, with one component for red, one for green and one for blue. In that case, the number of the output neurons is 3. For a greyscale image and same retina the number of the input neurons would be 9, while the number of the output neurons would be 1. Both, the number of layers and the number of elements in each layer have impact on the efficiency of the neural network. The program offers an option to choose the number of hidden layers, as well as the parameter of value for each hidden layer which shows how many times is the number of neurons in the hidden layer larger than the number of neurons in the input layer.
Each neuron consists of four basic functions: receipt of input, input processing, converting of the processed input into the output and synapse connections with other neurons. Network inputs are represented with the mathematical symbol \( x(n) \), and each input information is multiplied by its weight \( w(n) \). The results are added up (sum function):

\[
Sum = \sum w_i x_i
\]

and forwarded to the activation function which generates the result and then forwards it as the output.

### 3.3 Activation function

The activation function is used to scale the output data from the layers. The reason for using the activation function is the possibility for the neural network to learn a nonlinear function. If this function did not exist, the neurons from the hidden layers would not allow for more possibilities than the ordinary network that only consists of the input and the output – the perceptron.

For the activation function a sigmoid function has been chosen. The reason for using this function is the expected positive value as the output.

\[
f(x) = \frac{1}{1+e^{-x}}
\]

The output of the sigmoid function is the value between 0 and 1.

### 3.4 Training algorithm

Resilient propagation has been chosen for the learning algorithm, as suggested by the authors of the Encog3 neural networks. The advantage of this algorithm is that it does not require setting of the parameters such as the learning rate or the momentum value. That is good because it is sometimes very hard to determine these parameters in search of an optimal solution. The algorithm can be easily swapped in the program, but the examples tested and presented here use the resilient propagation algorithm.

Training is the way in which neural networks learn how to improve the weights of synaptic connections in order to achieve the desired result. A training set of input data and the ideal output for each input is forwarded to the training algorithm. The resilient propagation training algorithm goes through series of iterations and after each one it attempts to lower the error rate for a certain degree. The error rate represents a percent of difference between the real output and the output produced by the training algorithm. A gradient of error is calculated for each connection in the neural network. There is no global update parameter; rather for each value of the weight matrix there is a separate delta value. These values are at first randomly initiated at very low values. They are updated in accordance with delta values after each iteration. The magnitude of the gradient is used to determine how delta values should change so that each weight matrix can be individually trained [12].

The program used for experiments is written in JRuby programming language, version 1.9. JRuby is a Java implementation of the Ruby programming language, which enables usage of all the Java libraries by writing the Ruby code. It is a dynamic and object orientated programming language.

The proposed software uses Encog3 neural networks. Encog3 is an advanced framework for neural networks and machine learning. It contains classes for creating different types of networks, and also supports classes for normalizing and processing data for a specific type of networks. Encog3 propagates multithreading programming, works in many integrated developing frameworks, such as Eclipse or Netbeans. It is distributed as a JAR file. The software is developed in Netbeans programing framework. Resulting filter is saved as a neural network in the .eg format. The software supports different image formats such as .jpg, .gif, .png.

### 4 Experiments

In this section the results achieved by using the proposed method with our software are presented. For each selected filter the result of learning it will be shown first, followed by the results of application of that filter to newly selected images.

The simplest filters are the operations on one pixel. They are defined as a function carried out on each pixel of the image, independently from the other pixels of that same image. Some of the point processes are: converting a color image into a greyscale one, enlargement and reduction of contrast or brightness, the image negative and the threshold. As examples of the point process two filters are selected: the negative and the threshold.

The negative filter is particularly suited for improving the white and grey details surrounded by dark regions of the image, especially when black surfaces predominate. Figure 1 shows the result of learning the “invert” filter.

The obtained result is shown in the apprentice window. The error of the obtained result is 0.01, and the number of iterations is 965. Size of the retina was 3x3.
Fig. 1 Learning of the filter “negative”, upper row: original and the negative, lower row: “artificial negative” by our filter

Figure 2 illustrates the application of our “negative” filter to different image.

Fig. 2 Application of our filter “negative” to different image

An interesting example is double application of our “negative” filter which should reverse back to the original image (Figure 3)

Fig. 3 Application of our “negative” two times

It is remarkable how good result is considering that our filter does not “know” anything about concept of “negative”; it is a result of neural network learning about some pixels.

The main goal of image detection is deriving the important information from the image which can enable us to perform the computer interpretation and image analysis. The edges correspond to significant changes of image intensity. Intuitively, the edge is a set of connected pixels located on the border between two regions. There are many filters for edge detection and some of them are: Roberts cross, Sobel, Prewitt, Kirsh, Canny and the Laplacian filter. Figure 4 shows learning of the edge detection filter.

Fig. 4 Learning of the filter “edge”

The obtained result is shown in the apprentice window. The error of the result is 0.01, and the number of iterations is 12602. The retina used is 5x5.

Figure 5 shows the application of the learnt filter to newly selected image.

Fig. 5 Application of our filter “edge” to different image
The interpretation of the image depends on its sharpness, the possibility of isolating the information contained in the image. The sharper images show more detail. However, it is not easy to define what the sharp image is. It could be defined as an image that looks like a natural scene, but what appears natural to the human eye is difficult to calculate. Usually, a blurry image is defined as one that has lost the high frequency information. The images produced by using a photo camera are often blurry. Figure 6 shows learning of the deblur filter. Retina 5x5 was used. The training data has not been arbitrarily chosen; rather, the image surfaces were selected for the purpose:

The advantage of using our proposed method is the possibility to learn a filter which is a combined application to the image of larger number of filters. Image lena.png is taken as an example on which many filters have been applied: solarize, brightness, contrast and blur filter (Figure 8).

Fig. 7 Application of our filter “deblur” to different image

Fig. 6 Learning of the filter “deblur”

The result is shown in the apprentice window, the error is 0.506, and the number of iterations is 10320. Figure 7 illustrates the application of this filter to newly selected image.

Fig. 7 Application of our filter “deblur” to different image

The result is shown in the apprentice window. The error is 0.1, and the number of iterations is 9000. Figure 9 shows application of the learnt filter lena_mix.png to newly selected image:

Fig. 9 Application of our filter which replaces combination “solarize”, “brightness”, “contrast” and “blur” to different image

5 Conclusion
We developed a method based on image analogies and neural networks learning to determine a single filter that can reverse combined effect of multiple distortions. Even with only one pair of images for
training it is surprising how many different kinds and combinations of filters can be successfully determined by our method. The proposed solution for image analogy using neural networks has proved to be a good method for learning different kinds of filters, such as: threshold, conversion of a color image into a grey-scale one, image inversion, enlargement and reduction of contrast and brightness, detecting, blurring and sharpening of edges and also all kinds of their combinations. The method enables natural image transformations instead of the selection of different filters and their adjustments. The user can simply choose the desired effect and reproduce the it on new images. The proposed software also has its limitations: the filters that include larger distortions and those that depend on the context of the image itself.

References