Serial Number 09/111,370

Filing Date 30 June 1998

Inventor John M. Impagliazzo

NOTICE

The above identified patent application is available for licensing. Requests for information should be addressed to:

OFFICE OF NAVAL RESEARCH
DEPARTMENT OF THE NAVY
CODE 00CC
ARLINGTON VA 22217-5660

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

Navy Case No. 78695

CLASSIFICATION OF IMAGES USING A DICTIONARY OF COMPRESSED TIME-FREQUENCY ATOMS

STATEMENT OF GOVERNMENT INTEREST

The invention described herein may be manufactured and used by or for the Government of the United States of America for Governmental purposes without the payment of any royalties thereon or therefor.

BACKGROUND OF THE INVENTION

(1) Field of the Invention

The present invention relates to the field of image processing techniques, and more particularly to a method for automatically classifying test images based on their similarities with a dictionary of example target and non-target images organized according to class.

(2) Description of Related Art

The use of automatic pattern recognition systems and image classifiers for rapid identification and classification of input patterns (images) into one of several classes is well known in the art. Image classifiers have both military and civilian applications. For example, such systems can be used by a military combatant in a naval conflict to identify an unknown sonar target as a friend or foe, and thereby enable one to make an informed decision as to whether to attack the target. The systems are also used by civilians, for example, in medical screening and diagnostic applications. Additionally, image classification techniques are used for quality control in manufacturing applications.

Existing pattern recognition and image classification systems are typically based upon one of several conventional classification techniques. The conventional techniques for classifying images typically use a minimum set of manually distilled classification parameters from examples of known images which have been experimentally demonstrated to accurately classify a database of images into the correct class. For example, in the case of statistical classifiers, these parameters (features) consist of statistical moments scored according to a threshold criteria or nearest neighbor criteria. The features may also be based on ad hoc measurements or values defining properties of the image to be classified which have been proven successful on a test database. Additionally, classification parameters may be based on a model of the mechanisms which distinguish a class of images. Such conventional methods are well known in the art with examples being found in U.S. Patent Nos. 5,291,563 to Maeda, and 5,452,369 to Lionti et al.

In general, conventional automatic classifiers process a small set of clues derived from a large sequence of data representing the image to be classified. These conventional classification methods suffer from several significant drawbacks. One drawback is that the classification parameters or features used to classify an image are only a partial representation of the information in the image. Additionally, the methods are biased by the ad hoc algorithm used to quantitatively score the parameters used for classification. Furthermore, the existing techniques often are not easily modified for new or changing operational environments or when new input images or outcome classes are added. Often such changes require changing or modifying the features used for classification.

Accordingly, there is a need for a classification method which overcomes these drawbacks.

SUMMARY OF THE INVENTION

It is therefore an object of the present invention to provide a method of classification which operates by comparing a near-complete representation of a test image to a dictionary of example target and non-target images.

Another object of the present invention is the provision of a classification method which is easily augmented or refined for new operating environments.

The present method accomplishes these objects by receiving a test image and then initializing variables for an iteration count and for a linear expansion of the test image. The test image is then projected onto each one of the target and non-target images in the dictionary. A scaling coefficient is then applied for each successive iteration, wherein the scaling coefficient is set to the maximum value produced by the projections of the test image onto the dictionary of target and non-target example images. A residue is then generated, and the linear expansion of the test image is increased until a predetermined number of iterations have been performed.

Once this predetermined number of iterations have been performed, the sum of the scaling coefficients belonging to the target examples in the dictionary is compared to the sum of the scaling coefficients belonging to the non-target examples in the dictionary. If the sum of the scaling coefficients belonging to the target examples is greater than the sum of the scaling coefficients belonging to the non-target examples, then the test signal is identified as a target signal. If, however, the sum of the scaling coefficients belonging to the target examples is less than the sum of the scaling coefficients belonging to the non-target examples, then the test signal is identified as a non-target signal.

BRIEF DESCRIPTION OF THE FIGURES

A more complete understanding of the invention and many of the attendant advantages thereto will be readily appreciated and may be obtained from consideration of the following detailed description when considered in conjunction with the sole accompanying drawing which shows a flow diagram depicting an exemplary embodiment of the image classification technique according to the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The present invention discloses a method for automatically classifying two-dimensional test images based on their similarities with a dictionary of example images organized according to class. Like conventional image classification methods, the new method disclosed herein can be used for a variety of applications.

Images classified by the method of the present invention comprise two-dimensional arrays of pixels. Each pixel is assigned a value representing the gray level of that pixel. The pixel values can be distributed over any range. Additionally, the images can be comprised of more than one component array, such as color images. The image can be generated from an input signal using any conventional means such as digital cameras, scanners, acoustic imaging, an image previously stored in a digital format, or the like. In addition, the image can be processed as a whole or it can be divided into sub-images, with each sub-image being processed as a test image. Similarly, if a specific region of interest in the original image can be identified, the region can be processed as the test image.

This classification is accomplished by projecting a representation of the test image onto each of the example images in the dictionary. The projection process produces a representation of the test image as a linear expansion of scaled correlation coefficients in terms of the dictionary examples. The unknown image is then classified by comparing the scaling coefficients wherein if the sum of the scaling coefficients belonging to the target examples is greater than the sum of the scaling coefficients belonging to the non-target examples the unknown test image is identified as a target image, and otherwise the unknown test signal is identified as a non-target image.

The classification method disclosed herein employs an image compression technique which uses an invertible, lossy time-frequency transform. Although the images do not need to be compressed for the classification method of the present invention to operate. However, given that most images contain large numbers of pixels and that image processing is a computationally intensive procedure, the test images and the dictionary of images are usually compressed. An exemplary image compression algorithm which can be employed in connection with the method of the present invention is disclosed in U.S. Patent No. 5,757,974 to Impagliazzo *et al.* entitled *System and Method for Data Compression*. Other image compression techniques known in the art may also be used provided such methods maintain (preserve) a large majority of the original image information and can be reconstructed.

Images compressed utilizing the method of U.S. Patent No. 5,757,974 or the like contain a large majority of the original image information and can be readily reconstructed. Thus, when a compressed image is projected onto an example in a dictionary, all of the captured information is compared and the comparison is scored. As a result, this method is able to more accurately reproduce an unknown target image from the classification parameters than conventional methods.

The comparison is performed in the time-frequency domain because a substantial computational advantage is realized, equal to the compression ratio applied to the dictionary examples and test image. This is typically one to two orders of magnitude or larger. In addition, the score is a near complete representation of the test image. Also, since the dictionary consists of compressed time frequency transformed images, it can be augmented to include additional entries to refine the classifiers performance in other environments. This flexibility can be used to rapidly construct a classifier by developing a dictionary in a lab, on a test range, or in a similar controlled environment closely resembling an operational area of interest.

The present invention uses a matching pursuit algorithm disclosed in S.G. Mallat and Z. Zhang, *Matching Pursuit With Time-Frequency Dictionaries,* IEEE Trans. On Sig. Proc., Vol. 41, no. 12, pp. 3397-3415, December 1993, which allows a signal function to be decomposed into a linear expansion of functions belonging to a redundant dictionary of waveforms. In the present method, these waveforms are time-frequency atoms computed from both sample target and non-target images. It is assumed that the time-frequency atoms consist of a pattern of wavelet coefficients related to the local structure of the target. Without such an assumption, this information would otherwise be difficult to detect from individual coefficients because the forward transform diffuses the information across the entire basis. The present invention therefore employs an existing algorithm for a new purpose, i.e., image classification.

The advantage of the wavelet domain theory embodied in the method disclosed in the aforementioned article is that the respective image and dictionary waveforms can be compressed using wavelet image compression techniques, thereby preserving information about the local target structure without making any assumptions about the nature of the target. This compression, in turn, minimizes the computational requirements on the matching pursuit algorithm which defines a family of vectors $D=\left(g_\gamma\right)_{\gamma\in\Gamma}$ in H, where $H=L^2(R)$, such that $\|g_\gamma\|=1$. Letting $f\in H$, a linear expansion of f is computed over a set of vectors selected from D to best match the local target structure. This is done by successive approximations of f with orthogonal projections on elements of D. Letting $g_{\gamma_0}\in D$, the vector f can be decomposed into

$$f=\left\langle f,g_{\gamma_0}\right\rangle g_{\gamma_0}+Rf \tag{1}$$

where Rf is the residual vector after approximating f in the direction of g_{γ_0}. The element g_{γ_0} is orthogonal to Rf, hence

$$\|f\|^2=\left|\left\langle f,g_{\gamma_0}\right\rangle\right|+\|Rf\|^2. \tag{2}$$

To minimize $\|Rf\|$, $g_{\gamma_0}\in D$ is selected such that $\left|\left\langle f,g_{\gamma_0}\right\rangle\right|$ is maximized. To consider the iterative approach, let $R^0f=f$. To compute the n^{th} order residue $R^n f$, for $n \geq 0$, an element $g_{\gamma_n} \in D$ is chosen with the choice function C, which best matches the residue $R^n f$. The residue $R^n f$ is subdecomposed into

$$R^n f = \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^{n+1} f \qquad (3)$$

which defines the residue at the order n+1. Since $R^{n+1} f$ is orthogonal to g_{γ_n}

$$\left\| R^n f \right\|^2 = \left| R^n f, g_{\gamma_n} \right|^2 + \left\| R^{n+1} f \right\|^2 . \qquad (4)$$

Extending this decomposition to order m, equation (3) yields:

$$f = \sum_{n=0}^{m-1} \left\langle R^n f, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^m f \qquad (5)$$

and equation (4) yields an energy conservation equation:

$$\left\| f \right\|^2 = \sum \left| \left\langle R^n f, g_{\gamma_n} \right\rangle \right|^2 + \left\| R^m f \right\|^2 \qquad (6)$$

The original vector f is decomposed into a sum of dictionary elements that are chosen to best match its residues. Although the decomposition is nonlinear, it maintains an energy composition as if it was a linear orthogonal decomposition.

In utilizing the matching pursuit algorithm for target classification, the projection of the test image function onto each of the dictionary waveforms is computed. The waveform which best matches the image function is selected for the iteration and a residue is computed from the image function. The residue is formed by subtracting the selected waveform scaled by the correlation coefficient, from the image function to produce a new image function for the next iteration. After the last iteration, the image function is represented as a linear expansion of the scaled dictionary waveforms.

Target-like objects are discriminated from non-target image functions by comparing the energy in the dictionary's target waveform to that of the dictionary's non-target waveforms: The class associated with the greater energy is assigned to the image waveform. This process is shown in flowchart form in the sole figure in the case. The process starts as step 101 with the receipt of a test image x_i. At step 102, variables for the iteration count i and the linear expansion of the test image denoted by y are initialized, with i being set to 1 and y being set to 0.

At step 103, test image x_i is projected onto each of the images in the dictionary of example target and non-target images $D_{tgts+ntgts}$ and a scaling coefficient is identified. The scaling coefficient y_i for the i^{th} iteration is set to the maximum value produced by the projections of x_i onto the dictionary of images $D_{tgts+ntgts}$. The dictionary image which produces the maximum value when x_i is projected onto it, identified as $[D_{tgts+ntgts}]_i$, is associated with the scaling coefficient y_i. The projection of x_i onto the dictionary image is given as the inner product $\langle x_i, D_{tgts+ntgts} \rangle$ which produces a scalar quantity.

At step 104, the residue x_{i+1} is calculated by subtracting the dictionary image $[D_{tgts+ntgts}]_i$, identified in step 103 as producing the maximum result, scaled by y_i from x_i. That is, the residue x_{i+1} is given as:

$$x_{i+1} = x_i - y_i\left[D_{tgts+ntgts}\right]_i \qquad (7)$$

At step 105, the linear expansion of scaled dictionary waveforms y is refined by adding the scaled dictionary waveform y_i $[D_{tgts+ntgts}]_i$ to the existing linear expansion of scaled dictionary waveforms y. That is,

$$y = y + y_i\left[D_{tgts+ntgts}\right]_i \qquad (8)$$

The process of projecting x_i onto each waveform in the dictionary, generating the residue, and refining the linear expansion y is repeated until M iterations have been performed. If, as shown at step 106, fewer than M iterations have been performed, then at step 107 the number of iterations is incremented by 1 and the process is repeated from step 103. If however, M iterations have been performed, then at step 108 the sum of the scaling coefficients y_i belonging to the target examples in the dictionary $D_{tgts+ntgts}$ is compared to the sum of the scaling coefficients y_i belonging to the non-target examples in the dictionary. If the sum of the scaling coefficients y_i belonging to the target examples is greater than the sum of the scaling coefficients y_i belonging to the non-target examples, then the test signal is identified at step 109 as a target signal. If, however, the sum of the scaling coefficients belonging to the target examples is less than the sum of the scaling coefficients belonging to the non-target examples, then at step 110 the test signal is identified as a non-target signal.

The classification of target-like image functions is further refined by a back-propagation neural network. Such networks, which use artificial intelligence, are well known in the art and are used in the classification of test images. The neural network used need not be a back-propagation network but can be any type of neural network for classifying images. Although use of a neural network is not required to use the method of the present invention, such networks have been found to reduce the number of false alarms when classifying images. In using a neural network to further classify the test images, only the images identified as being targets are sent to the network. Because the input to the network is limited to those images which have been identified as targets, the construction of such a network is much simpler than that of a network that must distinguish targets from the set of all images. The input to the neural network can be the original image, a compressed image, a non-compressed image in the time frequency domain, or the linear expansion of scaled dictionary waveforms y.

In order to implement the matching pursuit/neural network classifier, it is necessary to divide the training set of data into subsets A and B. Half of the training set, subset A, is used as target waveforms for the matching pursuit dictionary. Non-target waveforms are also in the dictionary, but are selected from areas not proximate to the target. The remaining half of the training set, subset B, is processed using the matching pursuit algorithm having subset A in the target dictionary.

These results are then scored to form two lists for training the neural network: one of the functions for correctly classified targets, and one for the false alarms. The list of functions for correctly classified targets is augmented by an additional set generated from targets in training subset B with offset centers. The target and false alarm lists are used to train a neural network to discriminate targets from false alarms for the limited set of target-like image functions classified by the matching pursuit algorithm.

Successful application of this method is not limited to two-dimensional target images. The method described herein can be easily applied to classification of one dimensional signals or *n*-dimensional signals. In *n*-dimensional space, the compressed time-frequency representation of the signals is reshaped as a vector. The test signal vector is then projected onto equivalent signal vectors for each of the examples in the dictionary. In the one dimensional case, the compressed time-frequency representation of the signals is a vector. Typically, the number of iterations taken range between one an ten, but an alternative approach would be to increase the number of iterations further, but not to exceed the number of entries in the dictionary.

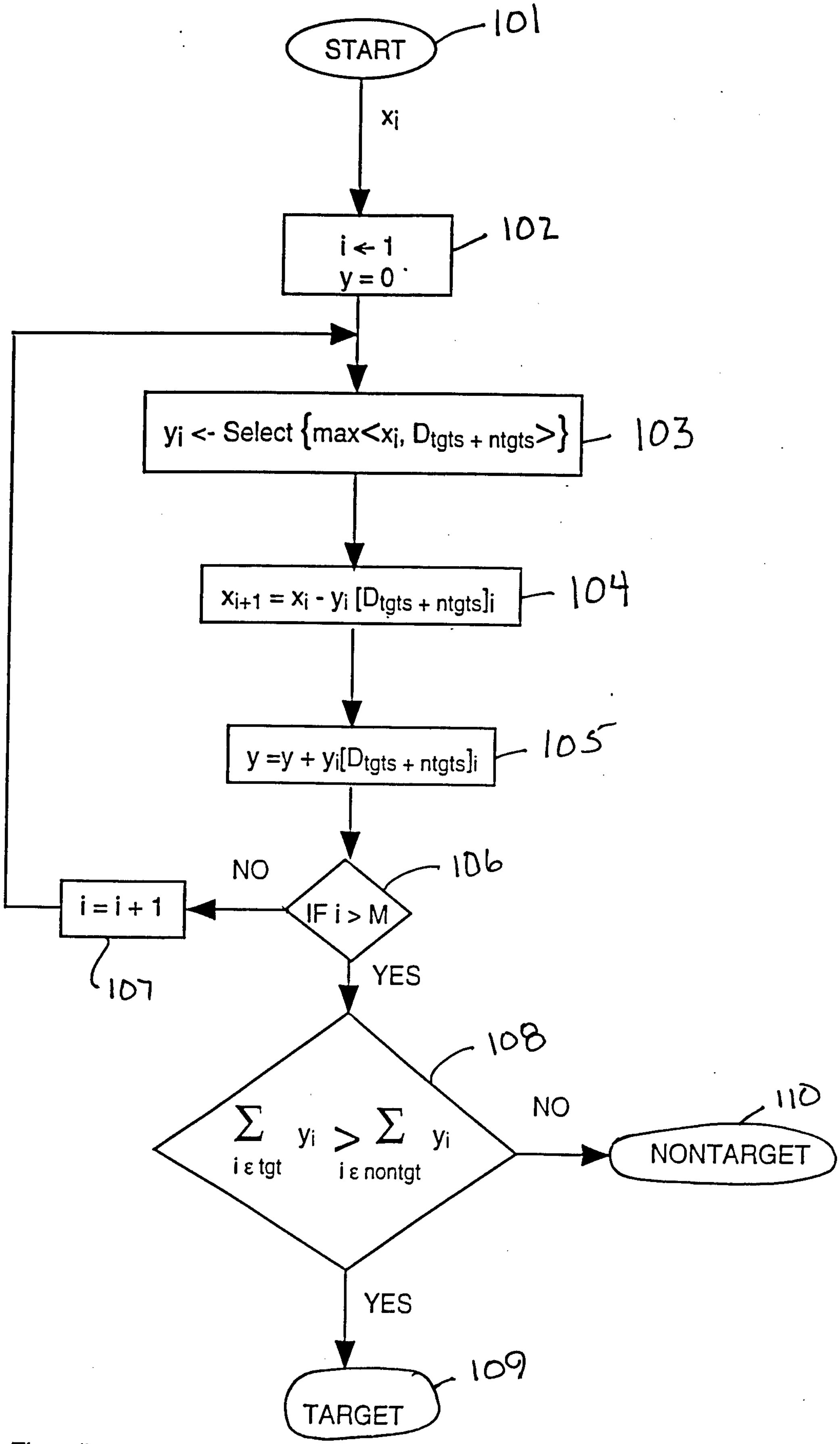
Numerous modifications to and alternative embodiments of the present invention will be apparent to those skilled in the art in view of the foregoing description. Accordingly, this description is to be construed as illustrative only and is for the purpose of teaching those skilled in the art the best mode of carrying out the invention. Details of the structure may be varied substantially without departing from the spirit of the invention.

Navy Case No. 78695

CLASSIFICATION OF IMAGES USING A DICTIONARY OF COMPRESSED TIME-FREQUENCY ATOMS

ABSTRACT OF THE DISCLOSURE

A method for automatically classifying test images based on their similarities with a dictionary of example target and non-target images. The method operates by receiving a test image and then initializing variables for an iteration count and for the linear expansion of the test image. The test image is then projected onto each one of the target and non-target images in the dictionary, wherein a maximum scaling coefficient is selected for each iteration. A residue is then generated, and the linear expansion of the test image is increased until a predetermined number of iterations have been performed. Once this predetermined number of iterations have been performed, the sum of the scaling coefficients belonging to the target examples in the dictionary is compared to the sum of the scaling coefficients belonging to the non-target examples in the dictionary to determine whether the image is a target signal or a non-target signal.



~~Figure 1.~~ ~~Flow diagram of Image classification technique~~