Title: METHOD AND DEVICE FOR CLASSIFYING SECURITY DOCUMENTS SUCH AS BANKNOTES

Abstract: A method for classifying security documents, such as banknotes, is provided. The method comprises a training and an operation phase. During the training phase at least one most discriminating feature among features of training security documents is determined. During the operation phase, a value of the at least one most discriminating feature is determined from the security document that are to be classified. The security document is classified on the basis of a comparison of the value with a threshold.
The invention relates to a method and a device for classifying security documents such as banknotes.

Many central banks are concerned with sorting security documents, in particular with determining whether banknotes are suitable for recirculation, or rather should be shredded and replaced by new ones. Obviously, more frequent recirculation reduces the printing costs and environmental burden. Given the huge amounts of different kinds of banknotes in circulation for even small countries, determining the fitness of banknotes poses a serious technical challenge in terms of processing speed and accuracy.

A device and a method for sorting banknotes is disclosed in applicant's European patent EP-B1-1043700.

It is an object of the present invention to provide an improved device and method for classifying security documents such as banknotes.
SUMMARY OF THE INVENTION

The object of the invention is met by providing a method for classifying security documents, such as banknotes, comprising the steps of:

During a training phase:
1) obtaining at least two digital training images of at least two respective training security documents;
2) dividing each digital training image into a set of predefined areas;
3) determining at least one feature for each area of each digital training image;
4) determining at least one most discriminating feature among all of the at least one feature and determining at least one threshold of the at least one most discriminating feature;

During an operating phase:
1) obtaining a digital image of a security document;
2) determining at least one value of the at least one respective most discriminating feature;
3) comparing the at least one value with the at least one respective threshold of the at least one most discriminating feature; and,
4) classifying the security document based on the comparison.

According to the invention classification of the security documents (or assigning each security document to a class) may take place on the basis of the most discriminating feature(s), while the most discriminating feature(s) is/are determined on the basis of a set of training security documents. Because of this, only a limited number (depending on how many most discriminating features are used in the operation phase) of features of the security documents that are to be classified, may need to be established and this may increase the speed of the classification process.
Since the features that are used to classify the security documents are determined as the most discriminating features for the set of training security documents, the accuracy of the classifying process during the operating phase may be high. The accuracy may be further increased by increasing the number of most discriminating features, as is explained below.

Furthermore, because of the determination of the most discriminating features, high accuracy in the classifying process may be obtained with a limited number of training security documents.

In an embodiment of the method according to the invention, step a2) further comprises: providing all areas of the set of predefined areas with respective first area weighting factors; and in step a4) said at least one most discriminating feature and said at least one threshold is determined using said first area weighting factors.

It may be advantageous to weigh the different areas when determining the at least one most discriminating feature and the at least one threshold. For example, one side of the security document or one specific area of that side may be more important or relevant with respect to classifying the security documents than others. For example, in classifying banknotes (an example of the security documents) the area comprising the text with the value of the banknote may be more relevant than the area comprising the signature of the bank president.

In an embodiment of the method according to the invention, step a3) further comprises: providing all features of said at least one feature with respective feature weighting factors; and, in step a4) said at least one most discriminating feature and said at least one threshold is determined using said feature weighting factors.

Since some features may be more relevant or important with respect to classifying the security documents than others, it may also be advantageous to weigh the different features.
In an embodiment of the method according to the invention, in step b4) the security document is classified based on the comparison using said first area weighting factors and/or said feature weighting factors.

It may be understood that weighting factors may be used for classifying the security documents, wherein the classification is based on the weighted combination of comparison of each of the at least one determined value with the at least one respective threshold.

It may be the case that the security document is classified as "fit" (or "unfit") when all of the at least one determined value are above their respective thresholds, when 50% (or another percentage) of all of the at least one determined value are above their respective thresholds, or when certain ones of the at least one determined value are above their respective thresholds.

In an embodiment of method according to the invention, wherein step a4) further comprises:
- providing all areas of the set of predefined areas with respective second area weighting factors;
- determining a first most discriminating feature using said second area weighting factors;
- adjusting said respective second area weighting factors; and,
- determining a second most discriminating feature using the adjusted second area weighting factors.

For example, before determining the n-th (n being a natural number > 2) most discriminating feature, the second area weighting factors may be adjusted, such that the second area weighting factors of those areas that are correctly classified by a combination of the n-1 most discriminating features may be decreased (or even be set equal to zero) and the second area weighting factors of those areas that are not correctly classified by a combination of the n-1 most discriminating features may be increased.
An advantage of adjusting the second weighting factors is that areas that have been incorrectly classified so far are more significant or important when the next most discriminating feature is determined.

In an embodiment of the method according to the invention, the predefined set of areas comprises overlapping rectangular areas with various sizes and aspect ratios. An advantage of this embodiment may be that it enables the determination of the same feature for different sized and overlapping areas. It may be the case that a feature is more discriminating when determined for an area A1 of the training security documents than when it is determined for an area A2 of the training security documents, wherein area A2 may contain area A1 completely. In this way, an optimum area for which the feature is most discriminating may be determined.

In an embodiment of the method according to the invention, the at least one feature comprises at least one of:

- an average of an intensity $I$ of said area; and,
- a standard deviation of the intensity $I$ of said area.

The intensity $I$ of an area of an image of a security document may be indicative of the state of the security document. For example, dirt on the security document may lower the intensity of an area of the image of the security document and/or the standard deviation of this intensity.

In an embodiment of the method according to the invention, the at least one feature comprises at least one of:

- an average of a red colour content $R$ of said area;
- a standard deviation of the red colour content $R$ of said area;
- an average of a blue colour content $B$ of said area;
- a standard deviation of the blue colour content $B$ of said area;
- an average of a green colour content $G$ of said area;
and,
- a standard deviation of the green colour content $G$ of said area;

Also the colour content of an area of an image of a security document may be indicative of the state of the security document, since certain kind of dirt may especially lower the reflection of certain colours.

In an embodiment of the method according to the invention, the at least one feature comprises at least one of:
- an average of a yellow-blue colour content $YB$ of said area;
- a standard deviation of the yellow-blue colour content $YB$ of said area;
- an average of a red-green colour content $RG$ of said area; and,
- a standard deviation of the red-green colour content $RG$ of said area;

wherein $YB$ is any linear combination of $R$, $G$ and $B$, such as $YB = R + G - 2B$, and $RG$ is any linear combination of $R$, $G$, $B$, such as $RG = R - 2G + B$.

Also the linear combination of colour contents of an area of an image of a security document may be indicative of the state of the security document. For example, dirt (or a sebum deposit) may be mainly apparent in the blue channel.

In an embodiment of the method according to the invention, the method comprises step b5) sorting the security document based on the classification.

In an embodiment of the method according to the invention, the at least one feature is normalized by an average of the intensity $I$ of a predetermined region of the digital training image. An advantage of this embodiment may be that variations in overall illumination during the scanning of the security documents may not influence the determination of a value of a feature.
In an embodiment of the method according to the invention, the intensity $I = R + G + B$.

In an embodiment of the method according to the invention, the step A1) comprises the steps of:

1) scanning the at least two respective training security documents to obtain at least two respective scanned images;

2) de-skewing and cutting said at least two respective scanned images to obtain at least two de-skewed scanned images; and,

3) fitting each of the at least two de-skewed scanned images into a predefined rectangular shape to obtain the at least two digital training images.

In an embodiment of the method according to the invention, the step B1) comprises the steps of:

1) scanning the security document to obtain a scanned image;

2) de-skewing and cutting said scanned image to obtain a de-skewed scanned image; and,

3) fitting each of the de-skewed scanned images into a predefined rectangular shape to obtain the digital image.

In a further embodiment of the method according to the invention, the step III) of A1) and/or step III) of B1) comprises aligning a security document printed image within said predefined rectangular shape.

The predefined rectangular shape of step III) of step A1) may correspond to the predefined rectangular shape of step III) of step B1).

When security documents such as banknotes are printed, a tolerance may be allowed in the exact positioning of the printed images relative to the paper boundary. Therefore, it may advantageous to align a security document printed image within said predefined rectangular shape in stead of aligning the paper boundary of the security document within the rectangular shape.

In an embodiment of the method according to the invention, the at least one most discriminating feature com-
prises at least 10 most discriminating features, or preferably 40 most discriminating features.

An advantage of using a higher number of most discriminating features may be an increased accuracy of the classification process. However, a higher number of most discriminating features may cause a longer processing time, during the training phase and/or during the operation phase. Furthermore, the increase of accuracy of a method using \( N+1 \) most discriminating features with respect to a method using \( N \) most discriminating features, decreases with a higher \( N \). Therefore, the number \( N \) of most discriminating features may be at its optimum at 40.

In an embodiment of the method according to the invention, the step b4) comprises classifying the security document as fit or as unfit.

In an embodiment of the method according to the invention, the step a2) further comprises determining the set of predefined areas on the basis of a contrast of the at least two digital training images.

When different kind of security documents are to be classified, the set of predefined areas may take into account all different layouts of all possible security documents. Using such a predefined set may cause the training phase to take a long time. It may therefore be advantageous to use a set of predefined areas that corresponds to the layout of the security documents to be tested.

On the basis of the contrast of the at least two digital training images information may acquired regarding the layout of the security documents to be tested. With this information, the predefined set of area may be determined. For example, the predefined set may be selected from a set of areas that takes into account all different layouts of all possible security documents.

The object of the invention is also met by providing a device for classifying security documents, such as banknotes, comprising:
- a training scanner arranged for obtaining at least two digital training images of at least two respective training security documents;
- a training image processing unit arranged for dividing each digital training image into a set of predefined areas and for determining at least one feature for each area of each digital training image;
- a training processing unit arranged for determining at least one most discriminating feature among all the at least one feature and at least one threshold of the at least one most discriminating feature;
- a scanner arranged for obtaining a digital image of a security document;
- an image processing unit arranged for determining at least one value of the at least one respective most discriminating feature;
- a classifying unit arranged for comparing the at least one value with the at least one respective threshold and for classifying the security document based on the comparison.

In an embodiment of the method according to the invention, the training scanner is the scanner, and/or the training image processing unit is the image processing unit.

In an embodiment of the method according to the invention, the device further comprises a sorting unit, arranged for sorting the security document based on the classification.

The advantages of the embodiments of the device according to the invention may be similar or equal to the advantages of the embodiments of the method according to the invention, as is explained in this document.

The various aspects and features described and shown in the specification can be applied, individually, wherever possible. These individual aspects, in particular the aspects and features described in the attached dependent
claims, can be made subject of divisional patent applications.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will be elucidated on the basis of an exemplary embodiment shown in the attached drawings, in which:

Figure 1 schematically depicts the steps of the training phase of an embodiment of the method according to the invention;

Figure 2 schematically depicts the steps of the operation phase of an embodiment of the method according to the invention;

Figure 3 schematically shows an example of a scanned image;

Figure 4 depicts an example of a digital (training) image;

Figure 5 shows schematically an example of a division of the digital (training) image into a set of areas according to an embodiment of the invention; and,

Figure 6 schematically depicts a device for classifying security documents according to an embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

Classification of security documents such as banknotes may take place in order to determine whether they are fit or unfit for circulation in society. Other examples of security documents are licenses, such as a driver license, coupons and other document that are regularly exchanged among users and may represent a certain economic value.

The soiling of a security document may be the main reason for classifying a security document as unfit. Other aspects of fitness of a security document may be stains and limpness, which show a high correlation with the level
of soiling. In this document the classification of security documents may be particularly concerned with the classification of security documents with respect to their soiling.

For euro bank notes, it was concluded that the main soiling mechanism may be that fingerprint deposits cumulates and eventually forms a yellow/brownish layer of aged sebum. In addition, it may be the (gentle) touch of the human fingers causing soil (particularly sebum) adhesion on the elevated parts, the crumple- or fold lines, of the banknote, which may be revealing a structural yet in-homogeneous appearance.

Because of the in-homogeneous appearance of soil, it may be very difficult or impossible to describe the level of soiling of security documents by objective, quantized characteristics (or features) that are applicable for all kinds of security documents in order to classify them as fit or unfit.

In an embodiment of the invention, a number of most discriminating features are determined for a certain kind of security document during a training phase using a set of training security documents. This determination may take place by machine learning techniques. The most discriminating features are then used during an operation phase to classify security documents.

Classification may refer to assigning a security document to one or more of a set of classes. For example, the set of classes may consist of the class "fit" and the class "unfit" and in that case the classification refers to assigning a security document to the "fit" class or to the "unfit" class. But the set of classes may also comprise three or more classes, for example "fit", "unfit, but repairable" and "unfit and unrepairable".

A characteristic of a (training) security document may be determined on the basis of an image of said security document. In this document, features of the image or parts of the image of the security document are determined
rather than characteristics of the security document itself. It is assumed that the features of (parts of) the image represent characteristics of the security document. And thus that the features may relate to the level of soiling of the security document.

First, embodiments of the training phase of the method according to the invention are described below. In the training phase a number of most discriminating features is determined. Figure 1 schematically depicts the steps of the training phase of an embodiment of the method according to the invention.

A first step during the training phase may be obtaining at least two digital training images of at least two respective training security documents. Each of training security documents has been classified before the start of the training phase. For example, the training security documents may have been manually classified by a group of experts.

The number of training security documents used in the training phase may be around 150 in each of the classes, for example 150 security documents that are classified as "fit" and 150 security documents that are classified as "unfit".

Step 11, obtaining the digital training images of the training security documents, may comprise the steps of:
- step 15, scanning the at least two respective training security documents to obtain at least two respective scanned images;
- step 16, de-skewing and cutting said at least two respective scanned images to obtain at least two de-skewed scanned images; and,
- step 17, fitting each of the at least two de-skewed scanned images into a predefined rectangular shape to obtain the at least two digital training images.

In step 15 scanned images may be obtained for example by a (training) scanner. Figure 3 schematically shows an example of a scanned image 31. The image may be taken from
the front or the back side of the training security document, however at least two images of the same side of the training security documents are required in the training phase.

In step 16 the scanned imaged may be cut along the lines of the security paper area 32. This security paper area may be obtained using an intensity threshold above a noise level. To reduce noise, the red, green and blue colour content may be added together to form the intensity image \( I = R + G + B \). In this way, signal-to-noise ratio may be optimized and this may yield most of the paper region.

Then the scanned image may need to be de-skewed and/or fitted, allowing determination of a feature over similar regions of the security document. As such, a box may be fitted around the security document, from which skew parameters can be estimated and a linear transformation may map the pixels to a rectangular and fixed sized digital (training) image.

In an embodiment of the method according to the invention, the fitting of each of the de-skewed scanned image into a predefined rectangular shape to obtain the digital image comprises aligning a security document printed image within said predefined rectangular shape.

In general, a tolerance may be allowed in the exact positioning of the security document printed images (which may comprise offset and intaglio prints) relative to the paper boundary, when security documents are generated. To reduce the variation between regions introduced by the allowed tolerances in the printing process, it may be advantageous to align the security document printed images more accurate or precise. This may be achieved using the following steps.

From a selection of training security documents that are classified fit, the one security document inducing the least amount of variation when being overlaid on the other security documents of the selection, may be determined. For this, a security document with minimum summed absolute
colour difference between the pixels colour content of all other security documents may be taken, wherein the colour difference between two pixels may be considered to be the sum of the absolute differences between the three respective colour contents. The resulting image may yield the typical (or modal) positioning of the security document offset layers within said selection. This resulting image may be used as a reference image for alignment of all other security documents.

Alignment of a given security document image may then proceed as follows. The security document image is deskewed as described above. After that, the image may be shifted in an x- and y-direction within an n x n neighbourhood, and matched against the reference image for all possible shifts within the neighbourhood. The shift with minimum summed colour difference to the reference image may yield the best alignment between the given security document image and the reference image. In this way, security document images may be aligned to the major content of the printing layers, rather than to the security document paper area.

In the above the step of obtaining at least two digital training images of at least two respective training security documents during the training phase is described. It may be understood that similar steps and embodiments may also be applicable on the step of obtaining a digital image of a security document in the operation phase.

Figure 4 depicts an example of a digital (training) image 41, that may be used in an embodiment of the method according to the invention or by an embodiment of the device according to the invention.

In the next step during the training phase, step 12 in figure 1, each digital training image is divided into a set of predefined areas, wherein the predefined set of areas may comprise overlapping rectangular areas with various sizes and aspect ratios. A training image processing unit may be arranged for executing step 12.
Figure 5 shows schematically an example of a division of the digital (training) image 41 into a set of areas 51 according to an embodiment of the invention. Although in figure 5 areas 51 have a rectangular shape, areas 51 may have a circular or any other two-dimensional shape. Areas 51 may or may not overlap each other. Areas 51 may have various sizes and may cover a large portion of the digital training image, for example an area 51 may cover the whole of the digital training image. The area 51 may cover a security mark on the image, for example a depiction of the value that the document is representing or any other depiction.

The set of predefined areas may comprise a set of random areas, i.e. a set of areas with random sizes and random positions. The set of predefined areas may comprise areas that have been selected for a certain kind of security document. The set of predefined areas may comprise areas that are assumed to be suitable for all kinds of security documents.

In an embodiment of the invention, the set of predefined areas is determined on the basis of the digital training images during the training phase. On the basis of contrast or contrast patterns of the training images it may be determined which kind of security document is processed and a set of predefined areas may be selected accordingly.

A set of predefined areas may be generated, wherein the set comprises areas around regions with a high or a low contrast in comparison with other regions. Instead of contrast, also colour content patterns (of R, G and or B) of the training images may be used in determining or generating a set of predefined areas.

In the next step, step 13 of figure 1, at least one feature for each area of each digital training image is determined. A training image processing unit may be arranged for executing step 13. Examples of features are:
- an average of an intensity I of said area;
- a standard deviation of the intensity $I$ of said area, which may represent the contrast of the area;
- an average of a red colour content $R$ of said area;
- a standard deviation of the red colour content $R$ of said area;
- an average of a blue colour content $B$ of said area;
- a standard deviation of the blue colour content $B$ of said area;
- an average of a green colour content $G$ of said area;
- a standard deviation of the green colour content $G$ of said area;
- an average of a linear combination of $R$, $G$ and $B$, for example $YB = R + G - 2B$ or $RG = R - 2G + B$. An advantage of (each of these two linear combinations may be that they may form together with the intensity the three orthogonal axis in a three dimensional colour space, and they may thus decorrelate the two chromatic information channels and the intensity channel.

It may be the case that digital (training) images of (training) security documents are obtained using electromagnetic radiation with wavelength in the visible range or in the non-visible range. For example, using infrared (IR) or ultraviolet (UV) radiation. Therefore, in general, a feature may be an average or standard deviation of a colour content, in which the colour is defined by a wavelength range and this wavelength range may be in the visible spectrum, but may also be in the non-visible range, such as the IR or the UV range. And a feature may also be a (linear) combination of an average or a standard deviation of a colour content, in which the colour is defined by a wavelength range.

It may be the case that new security documents reflect more light in a white region of the security document, for example in a region with a watermark. Therefore, the light intensity (and its standard deviation) of such an area of a digital image of a new security document may be higher with respect to an old security document.
To counteract variations in overall illumination during the scanning of the security documents, for example due to accumulated dust on an image sensor of the scanner and variations in overall printing quality of the security document, any of the above listed features may normalized by an average intensity of a certain region or a colour content (for example R, G, and B, of that certain region). This certain region may be its respective area, or any other area or may be the whole digital image.

The use of the blue colour content B (average or standard deviation) may by advantageous, since a sebum deposit may be mainly apparent in the blue colour content.

In an embodiment of the invention, twelve features for each area 51 are determined during the training phase, being the average and the standard deviation of I, R, G, B, YB=R+G-2B and RG=R-2G+B. Furthermore, for each training security document, features from both the front and the back side of the training security document may be determined. This may result in a large set of features, i.e. (the number of areas on the front and back side) x (number of features, for example 12) . As the examples of YB and RG show, a feature may also be a combination (for instance a linear combination) of the features described above.

However, only a small number out of the set of features may be used during the operation phase.

Instead of providing rules on how to classify security documents on the basis of one or more of these features, it may be advantageous to apply machine learning techniques to determine which features are most discriminating.

In the next step, step 14 in figure 1, at least one most discriminating feature in the set of features (i.e. among all of the at least one feature) is determined and at least one threshold of these most discriminating feature is determined. A training processing unit may be arranged for executing step 14.
Using the features of the training security documents and the known classification of the training security documents, it may be established which of the features is the most discriminating. For example, a feature (for instance: the average of the blue colour content of a certain area 51, which is normalized by the average intensity of that area) may correctly classify 60% of the training security documents in a "fit" and an "unfit" class using a threshold of 0.4. The threshold may imply that a value below 0.4 corresponds to the fit class, while a value above 0.4 corresponds to the unfit class.

When no other feature in the set of features is better (i.e. correctly classifies a higher percentage of the training security documents), this feature may be identified as the first most discriminating feature. The next best feature may then be identified as the second most discriminating features with its threshold and so on. In this way, a number of most discriminating features and their respective thresholds may be determined.

It may be the case that each area, into which each digital training image is divided, is provided with a respective first area weighting factor and/or a respective second area weighting factor.

The first area weighting factors may be used during the training and the operating phase. The first area weighting factors indicate the areas which are important for classifying the security documents, for example an area which depicts the denomination of a banknote.

The second area weighting factors may be used only during the training phase. The second area weighting factors may be adjusted several times when the most discriminating features are determined.

For example, after the first most discriminating feature has been determined (or identified), the respective second area weighting factors may be adjusted.

The second area weighting factors of those areas that are correctly classified by the first most discriminating
feature may be decreased or may even be set equal to zero. The second area weighting factors of those areas that are not correctly classified by the first most discriminating feature may be increased.

The second most discriminating feature may then be determined with respect to all areas of all digital training image, said areas having an adjusted second area weighting factor.

It may be understood that the above may also be applied when determining (or identifying) the third, the fourth and following most discriminating features: before determining the n-th (n being a natural number > 2) most discriminating feature, the second area weighting factors may be adjusted, such that the second area weighting factors of those areas that are correctly classified by a combination of the n-1 most discriminating features may be decreased (or even be set equal to zero) and the second area weighting factors of those areas that are not correctly classified by a combination of the n-1 most discriminating features may be increased.

This may yield that, after selecting a first most discriminating feature, the areas of the training security documents will be reweighted such that already correctly classified areas become less important, whereas still incorrectly classified examples are emphasized in the selection of the next most discriminating feature.

The combination of most discriminating features may be a linear combination of most discriminating features.

Figure 2 schematically depicts the steps of the operation phase of an embodiment of the method according to the invention.

In step 21 a digital image of a security document that is to be classified, is obtained. The step 21 may be executed similar to step 11. Likewise, step 21 may comprise the above describes embodiments, for example the steps 15, 16 and 17. A scanner may be arranged for executing step 21.
In step 22 a value of the at least one respective most discriminating feature is determined. An image processing unit may be arranged for executing step 22. Following the example described above, the value of the average of the blue colour content of the certain area 51, which is normalized by the average intensity of that area, may be determined for the security document that is to be classified. This value may be 0.6 in this example.

In step 23, the (determined) value of the at least one most discriminating feature is compared with the at least one respective threshold of the at least one most discriminating feature (in the example the threshold is 0.4). And in step 24 the security document is classified on the basis of the comparison. In the example, the security document would be classified as unfit. A classifying unit may be arranged for executing steps 23 and 24.

In an embodiment of the invention, a step 25 is executed after step 24, wherein the security document is sorted based on the classification. In the example, the security document, which is classified as unfit, may be removed from circulation in the society. A sorting unit may be arranged for executing step 25.

For a most discriminating feature more than one threshold may be determined. For example, it may be the case that the security documents are to be classified in more than two classes, for example in three classes. In that case, a feature may have two thresholds. A value between 0 and 0.3 may correspond to a "fit" class, a value between 0.3-0.4 may correspond to an "unfit, but repairable" or "dubious" class and a value between 0.4-1.0 may correspond to an "unfit and unrepairable" class.

Figure 6 schematically depicts a device for classifying security documents according to an embodiment of the invention.

The device 61 for classifying security documents, such as banknotes may comprise:
- a training scanner 62 arranged for obtaining at least two digital training images of at least two respective training security documents,
- a training image processing unit 63 arranged for dividing each digital training image into a set of predefined areas and for determining at least one feature for each area of each digital training image;
- a training processing unit 64 arranged for determining at least one most discriminating feature among all the at least one feature and at least one threshold of the at least one most discriminating feature;
- a scanner 65 arranged for obtaining a digital image of a security document;
- an image processing unit 66 arranged for determining at least one value of the at least one respective most discriminating feature;
- a classifying unit 67 arranged for comparing the at least one value with the at least one respective threshold of the at least one most discriminating feature and for classifying the security document based on the comparison.

In an embodiment of the device, the device further comprises a sorting unit 68, arranged for sorting the security document based on the classification.

The training scanner 62 may be arranged for providing data regarding the at least two digital training images to the training image processing unit 63, which may be arranged for receiving said data. The training image processing unit 63 may be arranged for providing data regarding the set of predefined areas and the at least one feature for each area of each digital training image to the training processing unit 64. The training processing unit 64 may be arranged to receive said data.

The training processing unit 64 may be arranged for providing data regarding the at least one most discriminating feature and the at least one threshold to the classifying unit 67, which may be arranged to receive said data. The scanner 65 may be arranged for providing data
regarding the at least two digital images to the image processing unit 66, which may be arranged for receiving said data.

The image processing unit 66 may be arranged for providing data regarding the at least one value of the at least one most discriminating feature to the classifying unit 67, which may be arranged to receive said data. The classifying unit 67 may be arranged for providing data regarding the comparison to the sorting unit 68, which may be arranged to receive said data.

In an embodiment of the device, the training scanner 62 is the scanner 65; and/or the training image processing unit 63 is the image processing unit 66.

It is to be understood that the above description is included to illustrate the operation of the preferred embodiments and is not meant to limit the scope of the invention. From the above discussion, many variations will be apparent to one skilled in the art that would yet be encompassed by the spirit and scope of the present invention.
1. Method for classifying security documents, such as banknotes, comprising the steps of:
   during a training phase:
   a1) obtaining at least two digital training images of at least two respective training security documents;
   a2) dividing each digital training image into a set of predefined areas;
   a3) determining at least one feature for each area of each digital training image;
   a4) determining at least one most discriminating feature among all of the at least one feature and determining at least one threshold of the at least one most discriminating feature;

during an operating phase:
   b1) obtaining a digital image of a security document;
   b2) determining at least one value of the at least one respective most discriminating feature;
   b3) comparing the at least one value with the at least one respective threshold of the at least one most discriminating feature; and,
   b4) classifying the security document based on the comparison.

2. Method according to claim 1, wherein step a2) further comprises: providing all areas of the set of predefined areas with respective first area weighting factors;

3. Method according to claim 2 or 3, wherein in step b4) the security document is classified based on the comparison using said first area weighting factors.
4. Method according to any of claims 1-3, wherein step a4) further comprises:
- providing all areas of the set of predefined areas with respective second area weighting factors;
- determining a first most discriminating feature using said second area weighting factors;
- adjusting said respective second area weighting factors;
and,
- determining a second most discriminating feature using the adjusted second area weighting factors.

5. Method according to any of claims 1-4, wherein the predefined set of areas comprise overlapping rectangular areas with various sizes and aspect ratios.

6. Method according to any of claims 1-5, wherein the at least one feature comprises at least one of:
- an average of an intensity I of said area; and,
- a standard deviation of the intensity I of said area.

7. Method according to any of claims 1-6, wherein the at least one feature comprises at least one of:
- an average of a red colour content R of said area;
- a standard deviation of the red colour content R of said area;
- an average of a blue colour content B of said area;
- a standard deviation of the blue colour content B of said area;
- an average of a green colour content G of said area;
and,
- a standard deviation of the green colour content G of said area;

8. Method according to any of claims 1-7, wherein the at least one feature comprises at least one of:
- an average of a yellow-blue colour content YB of said area;
- a standard deviation of the yellow-blue colour content YB of said area;
- an average of a red-green colour content RG of said area; and,
- a standard deviation of the red-green colour content RG of said area;

wherein YB is any linear combination of R, G and B, such as YB = R + G - 2B, and RG is any linear combination of R, G, B, such as RG = R - 2G + B.

9. Method according to any of claims 1-8, further comprising step b5) sorting the security document based on the classification.

10. Method according to any of claims 6-9, wherein at least one of the at least one feature is normalized by an average of the intensity I of a predetermined region of the whole digital training image.

11. Method according to any of claims 1-10, wherein the intensity I = R + G + B.

12. Method according to any of claims 1-11, wherein step a1) comprises the steps of:

i) scanning the at least two respective training security documents to obtain at least two respective scanned images;

ii) de-skewing and cutting said at least two respective scanned images to obtain at least two de-skewed scanned images; and,

iii) fitting each of the at least two de-skewed scanned images into a predefined rectangular shape to obtain the at least two digital training images.

13. Method according to any of claims 1-12, wherein step b1) comprises the steps of:
i) scanning the security document to obtain a scanned image;

ii) de-skewing and cutting said scanned image to obtain a de-skewed scanned image; and,

iii) fitting the de-skewed scanned image into a predefined rectangular shape to obtain the digital image.

14. Method according to any of claims 12-13, wherein step iii) comprises aligning a security document printed image within said predefined rectangular shape.

15. Method according to any of claims 1-14, the at least one most discriminating feature comprises at least 10 most discriminating features, or preferably 40 most discriminating features.

16. Method according to any of claims 1-15, wherein step b4) comprises classifying the security document as fit or as unfit.

17. Method according to any of claims 1-16, wherein step a2) further comprises determining the set of predefined areas on the basis of a contrast of the at least two digital training images.

18. Device for classifying security documents, such as banknotes, comprising:
- a training scanner arranged for obtaining at least two digital training images of at least two respective training security documents;
- a training image processing unit arranged for dividing each digital training image into a set of predefined areas and for determining at least one feature for each area of each digital training image;
- a training processing unit arranged for determining at least one most discriminating feature among all the at
least one feature and at least one threshold of the at least one most discriminating feature;
- a scanner arranged for obtaining a digital image of a security document;
- an image processing unit arranged for determining at least one value of the at least one respective most discriminating feature;
- a classifying unit arranged for comparing the at least one value with the at least one respective threshold of the at least one most discriminating feature and for classifying the security document based on the comparison.

19. Device according to claim 18, wherein:
- the training scanner is the scanner; and/or
- the training image processing unit is the image processing unit.

20. Device according to any of claims 18-19, wherein the device further comprises a sorting unit, arranged for sorting the security document based on the classification.
INTERNATIONAL SEARCH REPORT

PCT/NL2012/050380

A. CLASSIFICATION OF SUBJECT MATTER

According to International Patent Classification (IPC) or both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G07D7/20

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>wo 2007/068867 AI (NCR CORP [US]; HE CHAO [GB]; ROSS GARY [GB]) 21 June 2007 (2007-06-21) page 3, line 1 - page 5, line 6 page 9, line 1 - page 10, line 4 page 19, line 8 - page 20, line 13 figures 1-4</td>
<td>1,5-11 , 15-20</td>
</tr>
</tbody>
</table>

[Special categories of cited documents]

*A* document defining the general state of the art which is not considered to be of particular relevance

*E* earlier application or patent but published on or after the international filing date

*L* document which may throw doubts on priority claim(s) one of which is cited to establish the publication date of another citation or other special reason (as specified)

*O* document referring to an oral disclosure, use, exhibition or other means

*P* document published prior to the international filing date but later than the priority date claimed

*"T"* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

*X* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is taken alone

*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

*A* document member of the same patent family

Date of the actual completion of the international search

17 August 2012

Date of mailing of the international search report

19/09/2012

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2

NL - 2280 HV Rijswijk

Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016

Paraf, Edouard

Form PCT/ISA/210 (second sheet) (April 2005)
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent document cited in search report</td>
<td>Publication date</td>
<td>Patent family member(s)</td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-----------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>WO 2007068867 Ai</td>
<td>21-06-2007</td>
<td>BR PI0619845 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR PI0619926 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR PI0620308 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BR PI0620625 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1964073 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1964074 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1964075 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1964076 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2009519532 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2009527027 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2009527028 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2009527029 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2007140551 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2007154078 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2007154079 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2007154099 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2007068867 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2007068923 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2007068928 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2007068930 Ai</td>
</tr>
<tr>
<td>US 2007031021 Ai</td>
<td>08-02-2007</td>
<td>AT 485574 T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1912932 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 101739749 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1752934 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 2189951 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2007073035 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2007031021 Ai</td>
</tr>
<tr>
<td>EP 1484719 A2</td>
<td>08-12-2004</td>
<td>EP 1484719 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2004247169 Ai</td>
</tr>
<tr>
<td>US 2011038012 Ai</td>
<td>17-02-2011</td>
<td>EP 2130363 A2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2011038012 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2008107525 A2</td>
</tr>
<tr>
<td>EP 1489562 Ai</td>
<td>22-12-2004</td>
<td>CN 1573828 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 101303779 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 101354808 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DE 60309299 T2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1489562 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 4349849 B2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 2005010873 A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2004260649 Ai</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 2011125641 Ai</td>
</tr>
</tbody>
</table>