FORM 2

THE PATENTS ACT, 1970

(39 of 1970)

&

THE PATENT RULES, 2003

COMPLETE SPECIFICATION

(See Section 10 and Rule 13)

Title of Invention:

SYSTEMS AND METHODS FOR TRAINING BASED DETECTION AND CLASSIFICATION OF APPAREL

Applicant:

Tata Consultancy Services Limited
A company Incorporated in India under the Companies Act, 1956
Having address:
Nirmal Building, 9th floor,
Nariman point, Mumbai 400021,
Maharashtra, India

The following specification particularly describes the invention and the manner in which it is to be performed

CROSS-REFERENCE TO RELATED APPLICATIONS AND PRIORITY

[0001] The present application claims priority from Indian Provisional Patent Application No. 1862/MUM/2015, filed on May 11, 2015, the entirety of which is hereby incorporated by reference.

TECHNICAL FIELD

[0002] The present disclosure relates to the field of apparel recognition and classification, and more particularly to smart style recognition and classification technique for consolidation of style images in the fashion industry.

BACKGROUND

[0003] Typically, in the fashion industry, apparel retailers and buyers are continuously working on new collections, for which they are required to click a lot of pictures from fashion shows or trade shows. Generally speaking, apparel retailers are required to visit several fashion shows as they are expected to remain abreast of latest collections and new styles. A fashion show may offer hundreds of new styles and collating them for further categorization can require significant time and effort.

[0004] Basically, collation requires recognition of apparel styles based on features such as outlines, prints, patterns, and the like. For example, an apparel may have a full sleeve, half sleeve, a solid print or a block print. Similarly, differentiation can be based on base category of the apparel such as skirt, trouser, shorts, coat, jacket, and the like.

[0005] The challenge therefore remains in consolidating apparel pictures in appropriate folders based on their styles for easy retrieval without requiring the designer to browse through bundles of pictures to pick relevant pictures. Buyers and designers see this as a major bottleneck as a lot of time and effort is consumed in collating and storing these pictures for quick reference in real time. Furthermore, this may involve duplication of work as there is no means for identifying similarly styled pictures that are already stored.

SUMMARY

[0006] This summary is provided to introduce concepts related to consumer engagement in a service industry such as financial services with products like insurance, mutual funds, stocks and the like. This summary is neither intended to identify essential features of the claimed subject matter nor is it intended for use in determining or limiting the scope of the present disclosure.

[0007] In accordance with the present disclosure, the primary objective is to provide systems and methods for apparel recognition and classification in accordance with recognized kind and type of garment. Another objective of the present disclosure is to provide systems and methods for apparel recognition and classification according to local contour based features, texture features and various pre-defined ratios of the garment. Yet another objective of the present disclosure is to provide methods and systems for proper classification and categorization of apparel images in appropriate folders that are made accessible for quick and easy retrieval.

[0008] In an aspect there is provided a computer implemented method for training an apparel classifier, the method comprising receiving a plurality of first level training images of at least one apparel; generating first level training data comprising (a) a chamfer distance matrix of sensitized chamfer metric for each of the received plurality of first level training images and (b) an n-dimensional vector comprising size and shape related parameters of the at least one apparel; performing a first level learned classification of the received plurality of first level training images based on the generated first level training data to generate first level classified images; receiving a plurality of second level training images of at least one apparel; generating second level training data comprising (a) Histograms of Oriented Gradients (HOG) features and (b) normalized Hue and Saturation histogram for each of the received plurality of second level positive training images; and performing a second level learned classification of the first level classified images based on the generated second level training data to generate second level classified images.

[0009] In an embodiment, the step of receiving a plurality of first level training images is followed by a step of pre-processing the at least one training image comprising enhancing the at least one training image; and segmenting the enhanced at least one training image to separate background from the foreground.

[0010] In an embodiment, the step of generating the first level training data comprises computing the chamfer distance matrix comprising elements computed by the chamfer metric; sensitizing the chamfer metric based on orientation information to obtain the sensitized chamfer metric, the orientation information comprising placement, position and orientation of edgels present in each of the received plurality of first level training images; detecting edges in each of the received plurality of first level training images; and generating the n-dimensional vector, wherein the parameters are based on angles and edge ratios corresponding to pre-defined edges from the detected edges.

[0011] In an embodiment, the step of sensitizing the chamfer metric comprises determining angles of the edgels in relation to a pre-defined common reference; computing an orientation difference as a minimum difference of angle between every two neighboring edgels; and adding a weighted sum of the orientation difference to the chamfer metric.

[0012] In an embodiment, the step of detecting edges comprises thresholding each of the received plurality of first level training images using Sobel edge detection to identify the edges; and detecting continuity between the identified edges to form a straight line using Hough Detector.

[0013] In another aspect, there is provided a method for training based detection and classification of apparel comprising training an apparel classifier, wherein the step of training, as explained herein above can be followed by receiving at least one image of an apparel to be classified; computing the chamfer distance matrix of sensitized chamfer metric and the n-dimensional vector for the at least one received image; comparing the at least one received image with the first level classified images based on the computed chamfer distance matrix and the n-dimensional vector for the at least one received image to identify the at least one first level classified image; computing Histograms of Oriented Gradients (HOG) features and normalized Hue and Saturation histograms of the at least one first level classified image; and classifying the at least one first level classified image further by comparing with the second level classified images based on the computed HOG features and normalized Hue and

Saturation histograms for the first level classified image to identify the at least one second level classified image.

[0014] In an additional aspect, there is provided a system for training based detection and classification of apparel, the system comprising one or more processors; a communication interface device; one or more internal data storage devices operatively coupled to the one or more processors for storing an input module configured to receive a plurality of first level training images and second level training images of at least one apparel and at least one image of the apparel to be classified; a first level feature extractor configured to compute (a) a chamfer distance matrix of sensitized chamfer metric and (b) an n-dimensional vector comprising size and shape related parameters of the at least one apparel for each of the received plurality of first level training images and the at least one image of the apparel to be classified; a second level feature extractor configured to compute (a) Histograms of Oriented Gradients (HOG) features and (b) normalized Hue and Saturation histogram for each of the received plurality of first level training images and the at least one image of the apparel to be classified; a first level classifier configured to generate first level classified images pertaining to base category type of the apparel; and a second level classifier configured to generate second level classified images pertaining to texture of the apparel.

[0015] In yet another aspect, there is provided a computer program product comprising a non-transitory computer readable medium having program instructions embodied therein to perform a method to receive a plurality of first level training images of at least one apparel; generate a first level training data comprising (a) a chamfer distance matrix of sensitized chamfer metric for each of the received plurality of first level training images and (b) an n-dimensional vector comprising size and shape related parameters of the at least one apparel; perform a first level learned classification of the received plurality of first level training images based on the generated first level training data to generate first level classified images; receive a plurality of second level training images of at least one apparel; generating a second level training data comprising (a) Histograms of Oriented Gradients (HOG) features and (b) normalized Hue and Saturation histogram for each of the received plurality of second level positive training images; and perform a second level learned classification of the first level classified images based on the generated second level training data to generate second level classified images.

[0016] In an embodiment, the computer program product can be further configured to receive at least one image of an apparel to be classified; compute the chamfer distance matrix of sensitized chamfer metric and the n-dimensional vector for the at least one received image; compare the at least one received image with the first level classified images based on the computed chamfer distance matrix and the n-dimensional vector for the at least one received image to identify the at least one first level classified image; compute Histograms of Oriented Gradients (HOG) features and normalized Hue and Saturation histograms of the at least one first level classified image; classify the at least one first level classified image further by comparing with the second level classified images based on the computed HOG features and normalized Hue and Saturation histograms for the first level classified image to identify the at least one second level classified image; and organize the at least one first level classified image and the at least one second level classified image into labeled folders.

BRIEF DESCRIPTION OF DRAWINGS

[0017] The embodiments herein will be better understood from the following detailed description with reference to the drawings, in which:

[0018] FIG. 1 illustrates an exemplary block diagram of a system for apparel detection and classification in accordance with an embodiment of the present disclosure;

[0019] FIG. 2 illustrates a block diagram of the process flow in accordance with an embodiment of the present disclosure;

[0020] FIG. 3 illustrates a schematic representation of a first level learned classification in accordance with an embodiment of the present disclosure;

[0021] FIG. 4 illustrates a schematic representation of a second level learned classification in accordance with an embodiment of the present disclosure; and

[0022] FIG. 5 illustrates an exemplary garment with pre-defined edges that enable edge ratio and angle calculations in accordance with an embodiment of the present disclosure.

[0023] It should be appreciated by those skilled in the art that any block diagram herein represent conceptual views of illustrative systems embodying the principles of the present subject matter. Similarly, it will be appreciated that any flow charts, flow diagrams, state transition diagrams, pseudo code, and the like represent various processes which may be substantially represented in computer readable medium and so executed by a computing device or processor, whether or not such computing device or processor is explicitly shown.

DETAILED DESCRIPTION

[0024] The embodiments herein and the various features and advantageous details thereof are explained more fully with reference to the non-limiting embodiments that are illustrated in the accompanying drawings and detailed in the following description. The examples used herein are intended merely to facilitate an understanding of ways in which the embodiments herein may be practiced and to further enable those of skill in the art to practice the embodiments herein. Accordingly, the examples should not be construed as limiting the scope of the embodiments herein.

[0025] The words "comprising," "having," "containing," and "including," and other forms thereof, are intended to be equivalent in meaning and be open ended in that an item or items following any one of these words is not meant to be an exhaustive listing of such item or items, or meant to be limited to only the listed item or items.

[0026] It must also be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural references unless the context clearly dictates otherwise. Although any systems and methods similar or equivalent to those described herein can be used in the practice or testing of embodiments of the present disclosure, the preferred, systems and methods are now described.

[0027] Some embodiments of this disclosure, illustrating all its features, will now be discussed in detail. The disclosed embodiments are merely exemplary of the disclosure, which may be embodied in various forms.

[0028] Before setting forth the detailed explanation, however, it is noted that all of the discussion below, regardless of the particular implementation being described, is exemplary in nature, rather than limiting. For example, although selected aspects, features, or components of the implementations are depicted as being stored in memories, all or part of the systems and methods consistent with the apparel recognition and classification system and method may be stored on, distributed across, or read from other machine-readable media.

[0029] Broadly, the present disclosure aims at apparel recognition and classification that uses local contour based features, texture features and various critical ratios of the apparel, particularly garments for the purpose of detailed apparel recognition and classification.

[0030] Generally, the image of a garment is first recognized and classified according to its broad type, such as shirt, trousers, pants, skirts, top, etc. and subsequently, the classification can be based on its texture type. For example, a shirt may be striped, chequered, plain, floral, etc. Following this two stage classification, the image is categorized in pre-defined folders for the recognized kind and type of garment.

[0031] Expressions used interchangeably throughout the description include {garment, apparel, clothing}, {training. learning}, {contour, edge, line} and {cost, weight}.

[0032] The expression “thresholding” as referred to in the present disclosure pertains to dividing an image into ‘n’ number of sub-images for processing.

[0033] Referring now to the drawings, and more particularly to FIGS. 1 through 5, where similar reference characters denote corresponding features consistently throughout the figures, there are shown preferred embodiments and these embodiments are described in the context of the following exemplary system and method.

[0034] FIG. 1 illustrates an exemplary block diagram of system 100 for apparel detection and classification in accordance with an embodiment of the present disclosure. In an embodiment, system 100 includes one or more processors 10, communication interface or input/output (I/O) interface 12, and memory 14 or one or more internal data storage devices operatively coupled to one or more processors 10. One or more processors 10 can be implemented as one or more microprocessors, microcomputers, microcontrollers, digital signal processors, central processing units, state machines, logic circuitries, and/or any devices that manipulate signals based on operational instructions. Among other capabilities, Processor(s) 10 are configured to fetch and execute computer-readable instructions stored in memory 14. In an embodiment, system 100 can be implemented in a variety of computing systems, such as a laptop computer, a desktop computer, a notebook, a workstation, a mainframe computer, a server, a network server, cloud, hand-held device and the like.

[0035] I/O interface 12 can include a variety of software and hardware interfaces, for example, a web interface, a graphical user interface, and the like and can facilitate multiple communications within a wide variety of networks and protocol types, including wired networks, for example, LAN, cable, etc., and wireless networks, such as WLAN, cellular, or satellite. In an embodiment, the I/O interface can include one or more ports for connecting a number of devices to one another or to another server.

[0036] Memory 14 may include any non-transitory computer-readable medium known in the art including, for example, volatile memory, such as static random access memory (SRAM) and dynamic random access memory (DRAM), and/or non-volatile memory, such as read only memory (ROM), erasable programmable ROM, flash memories, hard disks, optical disks, and magnetic tapes. In an embodiment, modules 16 of system 100 can be stored in memory 14. Modules 16 can include routines, programs, objects, components, data structures, etc., which perform particular tasks or implement particular abstract data types to supplement applications and functions of system 100.

[0037] FIG. 2 illustrates a block diagram of the process flow or method 200, in accordance with an embodiment of the present disclosure, for apparel detection and classification. The present disclosure proposes learning of multilevel classifiers required for recognition and classification of different apparel images. Once the classifiers are learnt, apparel recognition and classification can be carried out. The process flow 200 of the present disclosure will now be explained with reference to the components of system 100 as depicted in FIG. 1.

[0038] The method of training the multilevel classifiers is detailed hereinafter. A plurality of first level training images of at least one apparel is received by input module 18. In an embodiment, first level training images 212 of base category types of apparel can include pants, shirts, skirts, jackets, dresses, blouses, kurtis, and the like. In an embodiment, first level training images 212 can be captured by system 100. First level training images 212 can comprise a plurality of first level positive training images and a plurality of first level negative training images. For instance, there can be 10 images of shirts as positive training images and 10 images of say pants as negative training images for classifying an apparel as a “shirt”.

[0039] In an embodiment, the step of receiving a plurality of first level training images 212 can be followed by a step of pre-processing (not shown) first level training images 212 by pre-processor 20. Preprocessing can involve enhancing features of first level training images 212 and reducing levels of noise using image-processing techniques to make the received first level training images 212 fit for passing on to a next stage. In an embodiment, preprocessing can include enhancing first level training images 212 using techniques known in the art such as Automatic Image Equalization and Contrast Enhancement Using Gaussian Mixture Modelling and separating the background from the foreground, using normalized cuts for performing image segmentation.

[0040] The step of receiving a plurality of first level training images 212 is followed by a multilevel classification stage for learning. This stage precedes the actual classification stage. The attempt is to train system 100 to perform the classification step at a later stage. In an embodiment, the process of training can be performed by feature extraction using contours, edge ratios and angles 214, by first level feature extractor 24 constituting feature extractor 22 and feature extraction using Histograms of Oriented Gradients (HOG) features 228, by second level feature extractor 26 constituting feature extractor 22.

[0041] In an embodiment, first level feature extractor 24 is configured to generate a first level training data that includes a chamfer distance matrix of sensitized chamfer metric for each of the received plurality of first level training images 212 and an n-dimensional vector comprising size and shape related parameters of the at least one apparel, wherein ‘n’ represents the number of sub-images that are formed by thresholding during processing of the images. FIG. 3 illustrates a schematic representation of first level learned classification 210 (of FIG. 2) in accordance with an embodiment of the present disclosure, wherein at a very broad level, canny edge detection 312 and oriented chamfer matching 314 is performed on input image 310 to generate a chamfer distance matrix; and Sobel and Hough edge detection 316, edge ratio and angle calculation 318 is performed on input image 310 to generate an n-dimensional vector comprising size and shape related parameters of the at least one apparel, the chamfer distance matrix and the n-dimensional vector constituting the first level training data provided to first level classifier 30 to generate first level classified images 222. In accordance with the present disclosure, Sobel edge detection comprises thresholding of images into ‘n’ number of sub-images to identify the edges.

[0042] Chamfer distance is a value of dis-similarity between two images say a query image and a learned image. Chamfer matching algorithm basically calculates the distance (dis-similarity) between the query image and the learned image. The basic idea is to extract edges or contours of the query image, take one point or pixel of contours in the query image and find the distance of closest points or pixels of contours in the learned image and sum the distances for all edge points or pixels of the query image. In an embodiment, the edges can be extracted by canny edge detection method. The lower the value of chamfer distance, the better the result. However the chamfer distance matrix is not very sensitive to the orientation of edgels or edge elements forming the edges or contours in images. The chamfer metric that computes the elements of the chamfer distance matrix is therefore sensitized by adding orientation information to perform oriented chamfer matching or to obtain the sensitized chamfer metric. The orientation information can include placement and position of edgels present in each of the received plurality of first level training images.

[0043] In an embodiment, the orientation information is added to the chamfer metric by first determining the angles of the edgels present in an image in relation to a common reference such as a line passing through the centroid of the image. The minimum difference of the angles of two neighboring edgels is then computed. This process is repeated for all the edgels present in the image. Once this has been done, a weighted sum of this orientation difference is added to the existing chamfer metric. The result is a metric which primarily depends on two factors: the placement and position of the edgels present in the image (i.e. the classic chamfer metric) and the relative orientations of the edgels of the image (sensitizing). Thus, the sensitized chamfer metric includes an explicit cost for the mean difference of orientations of the images. The chamfer metric can be presented by a relationship as given below –

$$d_{\text{orient}}^{(T,E)}(\mathbf{x}) = \frac{1}{N_T} \sum_{\mathbf{t} \in T} |o(\mathbf{t}) - o(\text{ADT}_E(\mathbf{t} + \mathbf{x}))|$$

wherein o (·) is the orientation of an edgel,

|o1 −o2| gives the smallest difference in angle (modulo π),

N_T is the Number of contour Fragments

T or Template is the contour fragment

t is the member of set T such that T = {t}

x is the relative position or the sub-image,

d is the chamfer score of x,

E is the edge map of an image,

ADT stands for argument distance transform

The argument distance transform, represented as,

$\mathrm{ADT}_E(\bar{\mathbf{q}}) = \arg\min_{\mathbf{e}\in E} \|\mathbf{q} - \bar{\mathbf{e}}\|_2$, wherein q is the argument of the ADT funtion and e is the member of set E such that E = {e}, can be computed simultaneously with the chamfer metric.

In accordance with the present disclosure, the chamfer distance measure is then sensitized by including the weighted sum between the chamfer distance and orientation terms:

$$d^{(T,E,\lambda)}(\mathbf{x}) = d^{(T,E)}_{\mathrm{cham}_\tau}(\mathbf{x}) + \lambda d^{(T,E)}_{\mathrm{orient}}(\mathbf{x})$$

where λ is an orientation specificity parameter. By learning a separate λ for each feature, a fine-grained control of orientation sensitivity is gained that could not have be obtained by dividing the edgels of E and T into discrete orientation channels.

[0044] In accordance with an embodiment, the step of sensitizing the chamfer metric is followed by the step of detecting edges in each of the received plurality of first level training images. In an embodiment, the step of detecting edges can include thresholding each of the received plurality of first level training images using Sobel edge detection to identify the edges and detecting continuity between the identified edges to form a straight line using Hough Detector. In an embodiment, the step of detecting edges or lines can be represented as an algorithm provided below –

Algorithm 1 Line Detection Algorithm for adjustment of parameters using Hough and Sobel Edge Detection

```
procedure GETLINES(sobelThresh, houghThresh, img)
    maxLines ← (img.ht/10)
    while linDtcd > maxLines do
        linDtcd ← HOUGHLINES(sobelThresh, houghThresh, img)
        incr ← (linDtcd/maxLines)
        houghThresh ← (houghThresh + incr)
        sobelThresh ← sobelThresh + (incr/2)
    end while
    Lines ← houghTransform(sobelThresh, houghThresh)
    FinalLines ← mean(Lines) returnFinalLines
end procedure
```

Every line in the image is detected multiple times. Firstly the method of detecting lines or edges in the image involves using arbitrary low values of thresholds for Sobel transform and Hough transform. In an embodiment, the low values can be less than 50. The threshold values for Sobel Transform is a grayscale value above which everything is considered as an edge (relevant for detection) and everything below this value is considered as background (irrelevant and can be ignored).

[0045] Every straight line in the image can have multiple edges, corresponding to it. For every edge, a pre-defined threshold value or a vote is assigned to the line to which it may belong. In the end, only those lines that have "votes" more than the pre-defined threshold value are detected as relevant. This process ensures that only prominent straight lines are detected. Accordingly, the threshold value for Hough Transform is the number of votes that enable a set of edges to be considered as a part of a straight line.

[0046] Once the number of lines in the entire image has been detected, using the arbitrary low values, a check is made on whether the total number of detected lines is less than at least a portion, for instance, a tenth (empirically determined) of the pixel height of the image. The empirically determined number of lines which maybe one tenth of the pixel height of the image is the maximum permissible number of lines in the image. In the event that the number of detected lines is more than at least a portion, for instance, a tenth of the pixel height of the image, an increment factor as the division of the number of lines detected and the maximum permissible number of lines is determined. The number of detected lines can be more than the pre-defined value generally at least in the first iteration. The Hough Threshold is then increased by the increment factor and the Sobel Threshold is increased by a half of the increment factor. The lines present in the image are recalculated using the new threshold values. The process is repeated till the number of detected lines are less than or equal to the maximum permissible number of lines. This process ensures detecting an optimum number of lines in the image.

[0047] In an embodiment, the parameters of the n-dimensional vector constituting the first level training data are based on angles and edge ratios corresponding to pre-defined edges from the detected edges as described herein above. The edge ratio and angle calculations ensure that the critical parameters of the detected garments are within predefined range of the detected base category of garments, for instance "T-Shirts". It would also be useful for detection and separation of similar categories of apparel like Men's and Women's T-Shirts. In an embodiment, the edge ratio and angle calculation of the present disclosure involves thresholding the image using Otsu's method and then applying Probabilistic Hough transform on the thresholded image with a high number of votes, as a threshold value, to prevent noisy edgels from getting detected. In an embodiment the number of votes can be 100. FIG. 5 illustrates an exemplary garment with pre-defined edges that enable edge ratio and angle calculations in accordance with an embodiment of the present disclosure. An exemplary illustration of classifying Mens's and Women's T-Shirts based on pre-defined parameters (measured lengths / angle between pre-defined portions of a garment as illustrated) with pre-defined values as illustrated in FIG. 5 is indicated in the table provided herein below.

Parameter	Men's T-Shirt	Women's T-Shirt
A1/A2	Larger	Smaller
B2/B1	Larger	Smaller
B2/B3	Larger	Smaller
B3/C	Larger	Smaller
θ	Larger	Smaller

[0048] In an embodiment, multilevel classifier 28 can include first level classifier 30 configured to generate first level classified images pertaining to base category type of the apparel and second level classifier 32 configured to generate second level classified images pertaining to texture of the apparel. For instance, first level classifier 30 can classify an apparel a base category type of a "shirt" and second level classifier 32 can further classify the "shirt" to be of type "plain" or "stripped" or say "chequered".

[0049] In an embodiment, first level learned classification 210 is based on the generated first level training data to provide first level classified images. The step of performing the first level learned classification 210 can include at least one of organizing the first level training data using a fragment dictionary with a large set of unlabeled images to improve the first level learned classification and enhancing the first level learned classification by boosting strong classifications and suppressing weak classifications. Fragment dictionary approach, as known in the art, is used to improve detection accuracy of first level classifier 30. Fragment dictionary is a collection of spatially localized contour fragments for a given image. These fragments are stored relative to the centroid of the image.

At the time of detection, the individual fragments are tried to be matched, instead of the entire image. Boosting is a standard process applied on each fragment of the fragment dictionary for improving the classification result. For instance, the process of boosting can identify whether the sensitized chamfer metric or the n-dimensional vector is stronger and weighted values may be accordingly adjusted.

[0050] In an embodiment, the step of performing the first level learned classification uses K-Nearest Neighbor (KNN) method. In an embodiment, the top priority features (chamfer metric) in the order of their importance for first level classifier 30, in this instance a KNN classifier, are:

Oriented Chamfer Match Distance: $d^{(T,E,\lambda)}(\mathbf{x})$ with an empirical weight of 0.3
Theta (θ) - the angle, with a weight of 0.3
B3/C - the edge ratio with a weight of 0.1
B2/B1 - the edge ratio with a weight of 0.1
B2/B3 - the edge ratio with a weight of 0.1
A1/A2 - the edge ratio with a weight of 0.1
It may be noted from the above weights that the sum of all the values adds up to one. In the exemplary embodiment, an image is classified into a particular base category, when the total weight of the positive features is greater than 0.5.

[0051] Once first level classifier 30 is trained, second level classifier 32 is accordingly trained. This will further classify the results of the first level classified garments having the same type (i.e. shirt or pant) into different classes based on the texture of the garment. For instance, a “shirt” as classified by first level classifier 30 can be further classified as “plain” or “stripped” or say “chequered” by second level classifier 32. First level classifier 30 and second level classifier 32 are required to be trained for every garment category. However when the styles are similar, the same classifier can be shared.

[0052] A plurality of second level training images of at least one apparel is received by input module 18. In an embodiment, second level training images 226 of different textures can include plain, stripped, chequered and the like. In an embodiment, second level training images 226 can be captured by system 100. Second level training images 226 can include a plurality of second level positive training images. For instance, there can be 10 images of plain shirts as positive training images for classifying a “shirt” as a “plain shirt”. Negative training images may be used depending on the method employed. For instance, Support Vector Machine (SVM) method for classification, as known in the art, would not benefit from use of negative training images.

[0053] The step of receiving a plurality of second level training images 226 by system 100 is followed by performing second level learned classification 224. In an embodiment, second level feature extractor 26 is configured to generate a second level training data that includes HOG features and normalized Hue and Saturation histogram for each of the received plurality of second level positive training images. Hue and Saturation histogram enables color based segmentation of the garments, especially of those garments that have a plain texture. FIG. 4 illustrates a schematic representation of second level learned classification 224 (of FIG. 2) in accordance with an embodiment of the present disclosure, wherein gray scale conversion 412, gradient detection 414 and HOG feature extraction 416 is performed on input image 410 to generate the second level training data for second level classifier 32. Gray scale conversion 412 involves converting a colored image to a black and white image. For instance, a standard 256 level RGB image is converted to an image having 256 different shades of gray varying from black (0) to white (255). The Histogram of Oriented Gradients (HOG) feature extraction involves detecting number of occurrences of gradient orientation in localized portions of an image. It involves detecting major orientation of all sub-images in a given image and is computed on a dense grid of uniformly spaced sub-images and uses overlapping local contrast normalization for improved accuracy.

[0054] In an embodiment, second level learned classification 214 is based on the generated second level training data to generate second level classified images 234.

[0055] Training of classifiers to facilitate classification of an input image of an apparel is a critical aspect of systems and methods of the present disclosure. Referring to FIG. 2, classification of an apparel can include the step of receiving at least one image of an apparel to be classified. Feature extraction using contours, edge ratios and angles 220 is performed to compute the chamfer distance matrix of sensitized chamfer metric and the n-dimensional vector for the received input image 218 in accordance with the steps described herein above in relation to training images. The received input image 218 is then compared with the first level classified images by first level classifier 30 based on the computed chamfer distance matrix and the n-dimensional vector to obtain first level classified images 222. Feature extraction using Histograms of Oriented Gradients (HOG) features 232 and normalized Hue and Saturation histograms of first level classified images 222 are computed in accordance with the steps described herein above in relation to training images. First level classified images 222 are then compared with the second level classified images by second level classifier 32 based on the computed HOG features and normalized Hue and Saturation histograms to obtain second level classified images 234.

[0056] In an embodiment, the step of performing a second level learned classification uses Support Vector Machine (SVM) method. Accordingly, there is no added benefit in using negative training images and therefore training images may be limited to positive training images.

[0057] The step of multilevel classification comprising first level learned classification 210 and second level learned classification 224 is followed by the step of organizing 236 the first level classified image 222 and second level classified image 234 into labeled folders by organizer 34 (of FIG.1) to facilitate identifying a pre-fed location where images will be located using system commands for easy retrieval. In an exemplary embodiment, the images can be moved into labeled folders in a LINUX environment using C. The database containing the folders can be made accessible on a cloud. In an embodiment, an automation command in Linux may be used for copying the image to desired folders.

[0058] It is to be understood that the disclosure is not limited in its application to the details of construction and parts illustrated in the accompanying drawings and described herein-above. The disclosure is capable of other embodiments and of being practiced in various ways. It is also to be understood that the phraseology or terminology used herein is for the purpose of description and not limitation. Hence, although the present disclosure has been described hereinabove by way of illustrative embodiments thereof, it can be modified at will, within the scope of the appended claims, without departing from the spirit, scope and nature of the subject disclosure.

WE CLAIM:

1. A computer implemented method for training an apparel classifier, said method comprising:

receiving a plurality of first level training images of at least one apparel;

generating first level training data comprising (a) a chamfer distance matrix of sensitized chamfer metric for each of the received plurality of first level training images and (b) an n-dimensional vector comprising size and shape related parameters of the at least one apparel;

performing a first level learned classification of the received plurality of first level training images based on the generated first level training data to generate first level classified images;

receiving a plurality of second level training images of at least one apparel;

generating second level training data comprising (a) Histograms of Oriented Gradients (HOG) features and (b) normalized Hue and Saturation histogram for each of the received plurality of second level positive training images; and

performing a second level learned classification of the first level classified images based on the generated second level training data to generate second level classified images.

2. The computer implemented method of claim 1, wherein the step of receiving a plurality of first level training images comprises capturing the plurality of first level training images of the apparel.

3. The computer implemented method of claim 1, wherein the step of receiving a plurality of first level training images is followed by a step of pre-processing the at least one training image comprising:

enhancing the at least one training image; and

segmenting the enhanced at least one training image to separate background from the foreground.

4. The computer implemented method of claim 1, wherein the step of generating the first level training data comprises:

computing the chamfer distance matrix comprising elements computed by the chamfer metric;

sensitizing the chamfer metric based on orientation information to obtain the sensitized chamfer metric, the orientation information comprising placement, position and orientation of edgels present in each of the received plurality of first level training images;

detecting edges in each of the received plurality of first level training images; and

generating the n-dimensional vector, wherein the parameters are based on angles and edge ratios corresponding to pre-defined edges from the detected edges.

5. The computer implemented method of claim 4, wherein the step of sensitizing the chamfer metric comprises:

(i) determining angles of the edgels in relation to a pre-defined common reference;

(ii) computing an orientation difference as a minimum difference of angle between every two neighboring edgels; and

(iii) adding a weighted sum of the orientation difference to the chamfer metric.

6. The computer implemented method of claim 4, wherein the step of detecting edges comprises:

(i) thresholding each of the received plurality of first level training images using Sobel edge detection to identify the edges; and

(ii) detecting continuity between the identified edges to form a straight line using Hough Detector.

7. The computer implemented method of claim 1, wherein the step of performing the first level learned classification uses K-Nearest Neighbour (KNN) method.

8. The computer implemented method of claim 1, wherein the step of performing a first level learned classification is followed by at least one of:

(i) organizing the first level training data using a fragment dictionary with a large set of unlabeled images to improve the first level learned classification; and

(ii) enhancing the first level learned classification by boosting strong classifications and suppressing weak classifications.

9. The computer implemented method of claim 1, wherein the step of performing a second level learned classification uses Support Vector Machine (SVM) method.

10. A computer implemented method for training based detection and classification of apparel comprising training an apparel classifier as claimed in claim 1, the step of training being followed by:

(i) receiving at least one image of an apparel to be classified;

(ii) computing the chamfer distance matrix of sensitized chamfer metric and the n-dimensional vector for the at least one received image;

(iii) comparing the at least one received image with the first level classified images based on the computed chamfer distance matrix and the n-dimensional vector for the at least one received image to identify the at least one first level classified image;

(iv) computing Histograms of Oriented Gradients (HOG) features and normalized Hue and Saturation histograms of the at least one first level classified image; and

(v) classifying the at least one first level classified image further by comparing with the second level classified images based on the computed HOG features and normalized Hue and Saturation histograms for the first level classified image to identify the at least one second level classified image.

11. The computer implemented method of claim 10, further comprising organizing the at least one first level classified image and the at least one second level classified image into labeled folders.

12. The computer implemented method of claim 10, wherein the step of receiving at least one image of an apparel to be classified is followed by a step of pre-processing the at least one image comprising:

enhancing the at least one image; and

segmenting the enhanced at least one image to separate background from the foreground.

13. The computer implemented method of claim 10, wherein the step of receiving at least one image of an apparel to be classified comprises capturing the at least one image of the apparel.

14. A system for training based detection and classification of apparel, said system comprising:

one or more processors;

a communication interface device;

one or more internal data storage devices operatively coupled to the one or more processors for storing:

an input module configured to receive a plurality of first level training images and second level training images of at least one apparel and at least one image of the apparel to be classified;

a first level feature extractor configured to compute (a) a chamfer distance matrix of sensitized chamfer metric and (b) an n-dimensional vector comprising size and shape related parameters of the at least one apparel for each of the received plurality of first level training images and the at least one image of the apparel to be classified;

a second level feature extractor configured to compute (a) Histograms of Oriented Gradients (HOG) features and (b) normalized Hue and Saturation histogram for each of the received plurality of first level training images and the at least one image of the apparel to be classified;

a first level classifier configured to generate first level classified images pertaining to base category type of the apparel; and

a second level classifier configured to generate second level classified images pertaining to texture of the apparel.

15. The system of claim 14 further comprising an organizer configured to locate the first level classified images and the second level classified images into labeled folders.

16. The system of claim 14, further comprising a pre-processor configured to enhance the at least one training image and image to be classified; and segment the enhanced at least one training image and the image to be classified, to separate background from the foreground.

17. The system of claim 14, wherein the plurality of first level training images comprise a plurality of first level positive training images and a plurality of first level negative training images.

18. The system of claim 14, wherein the plurality of second level training images comprise a plurality of second level positive training images.

Dated this 27th day of October, 2015

Amit Koshal
IN-PA-2358
Agent for the Applicant

ABSTRACT

SYSTEMS AND METHODS FOR TRAINING BASED DETECTION AND CLASSIFICATION OF APPAREL

The present disclosure envisages a computer implemented system and method for apparel recognition and classification that uses local contour-based features, texture features and various pre-defined ratios of the garments for the purpose of detailed apparel recognition and classification. The system makes use of a two stage process including partially supervised learning architecture and a multilevel classifier for learning the garment image features. The system performs steps of determining pre-defined features of images of the garment received for classification, performing multilevel classification of the image using the pre-learned classifier, recognizing the image based on the result of learned features, followed by proper classification of the image and organizing the images in labeled folders.

[*To be published with Figure 2*]

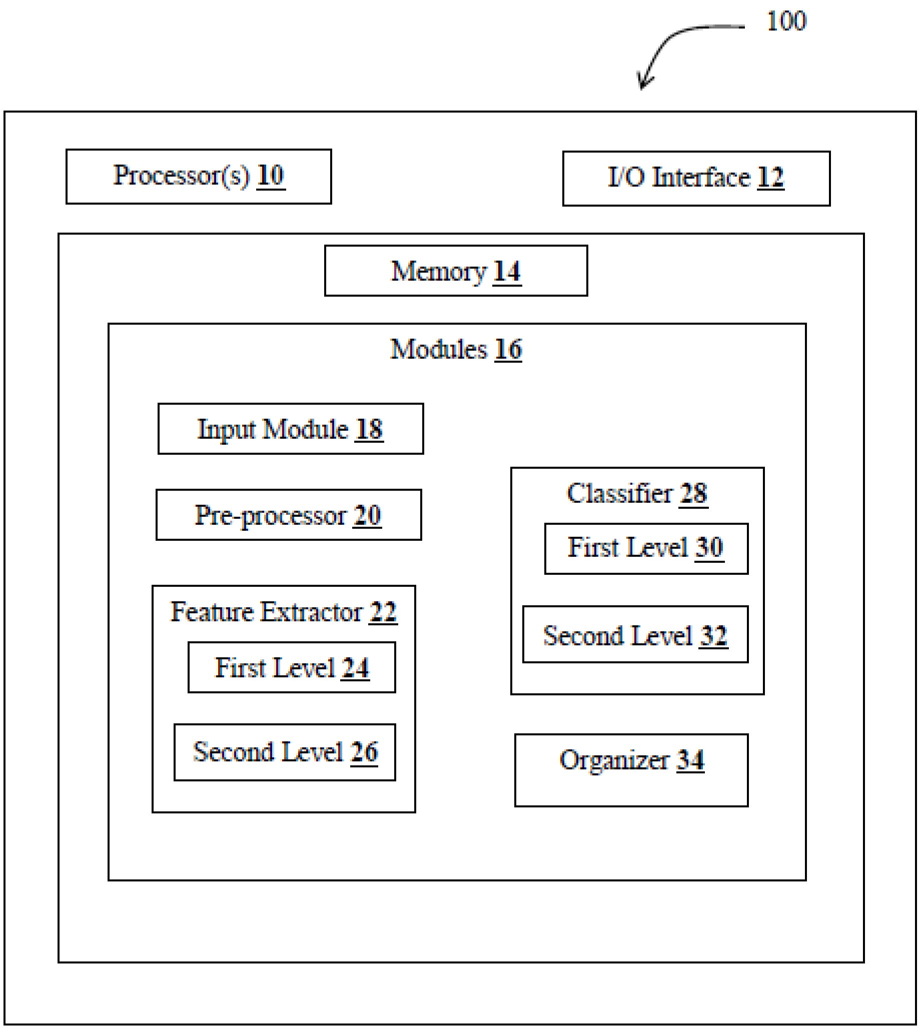

FIG. 1

Amit Koshal
IN-PA-2358
Agent for the Applicant

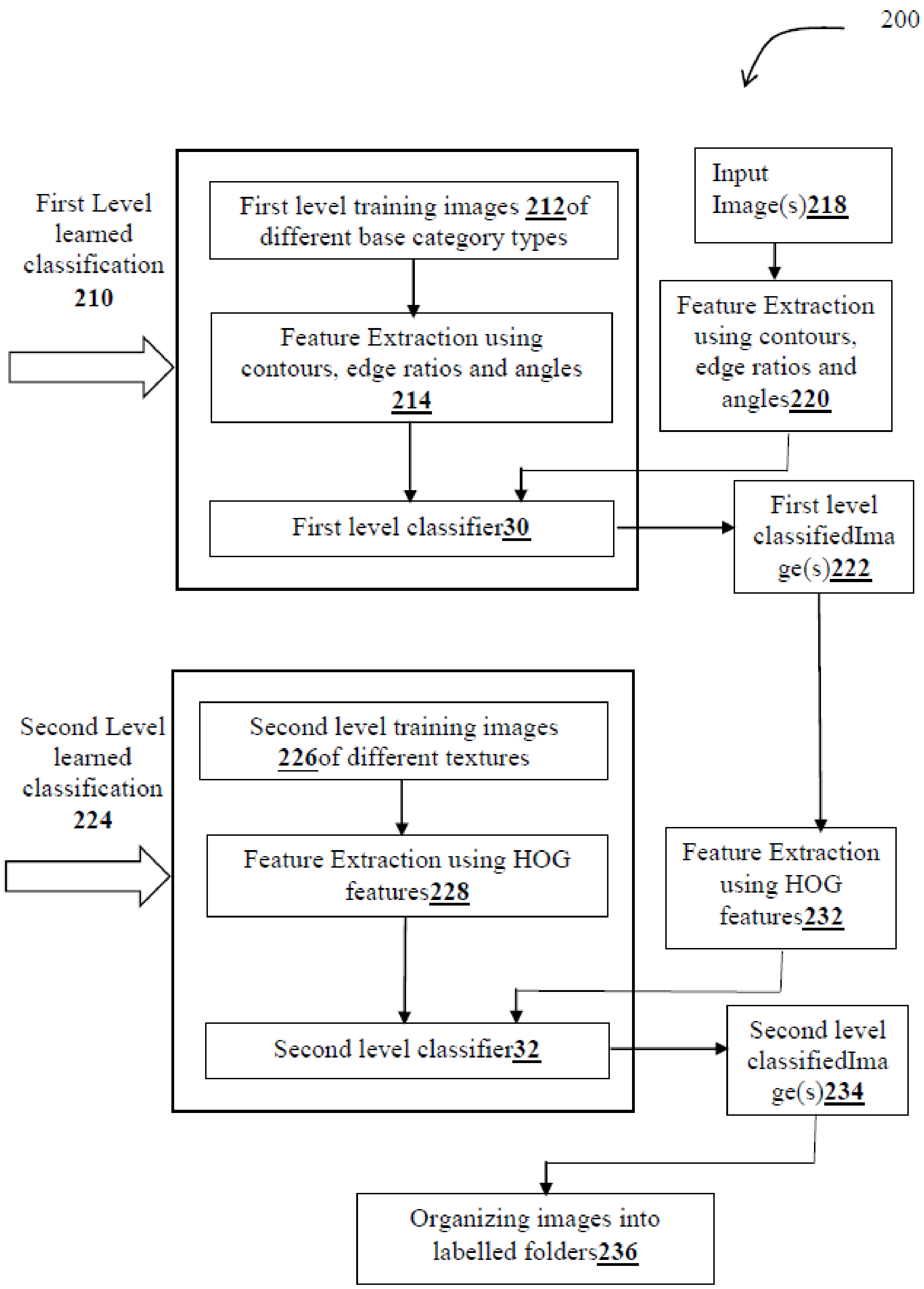

FIG.2

Amit Koshal
IN-PA-2358
Agent for the Applicant

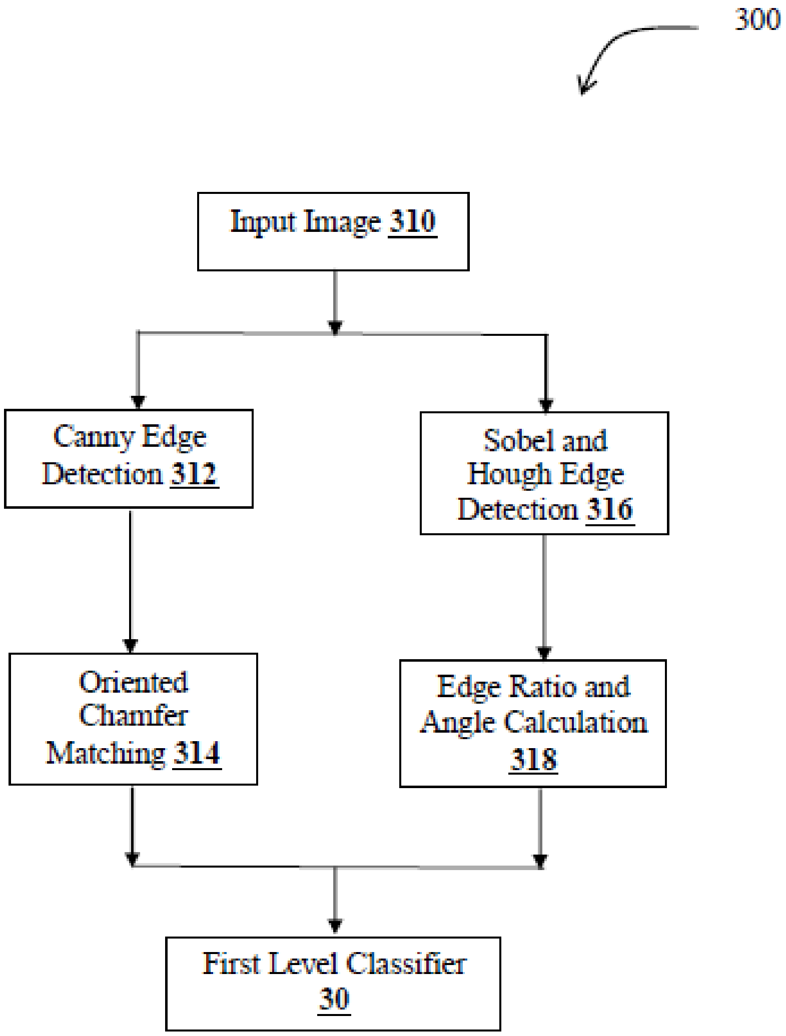

FIG.3

Amit Koshal
IN-PA-2358
Agent for the Applicant

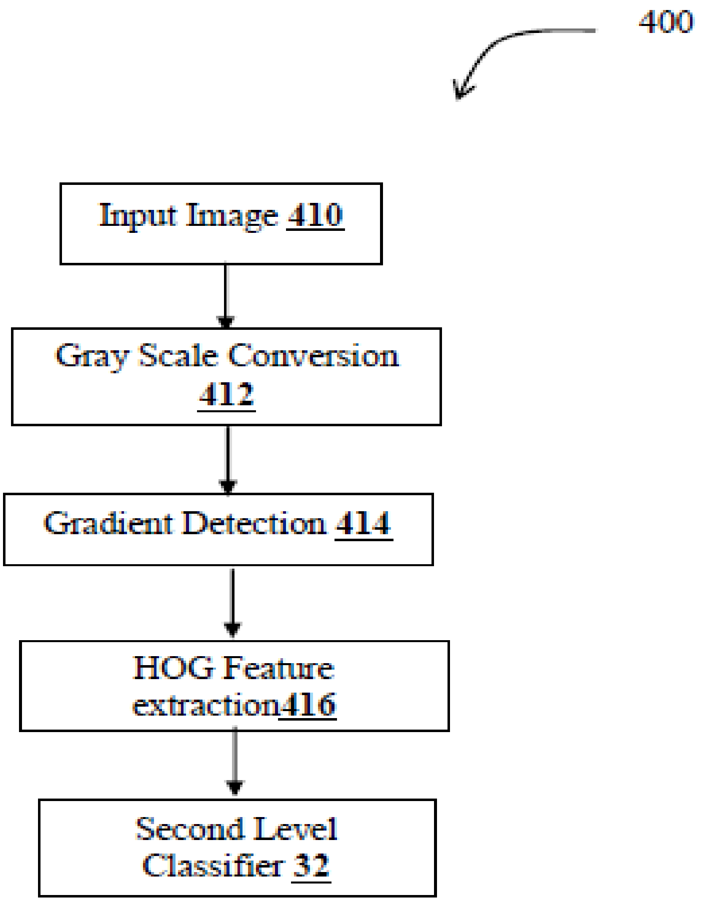

FIG.4

Amit Koshal
IN-PA-2358
Agent for the Applicant

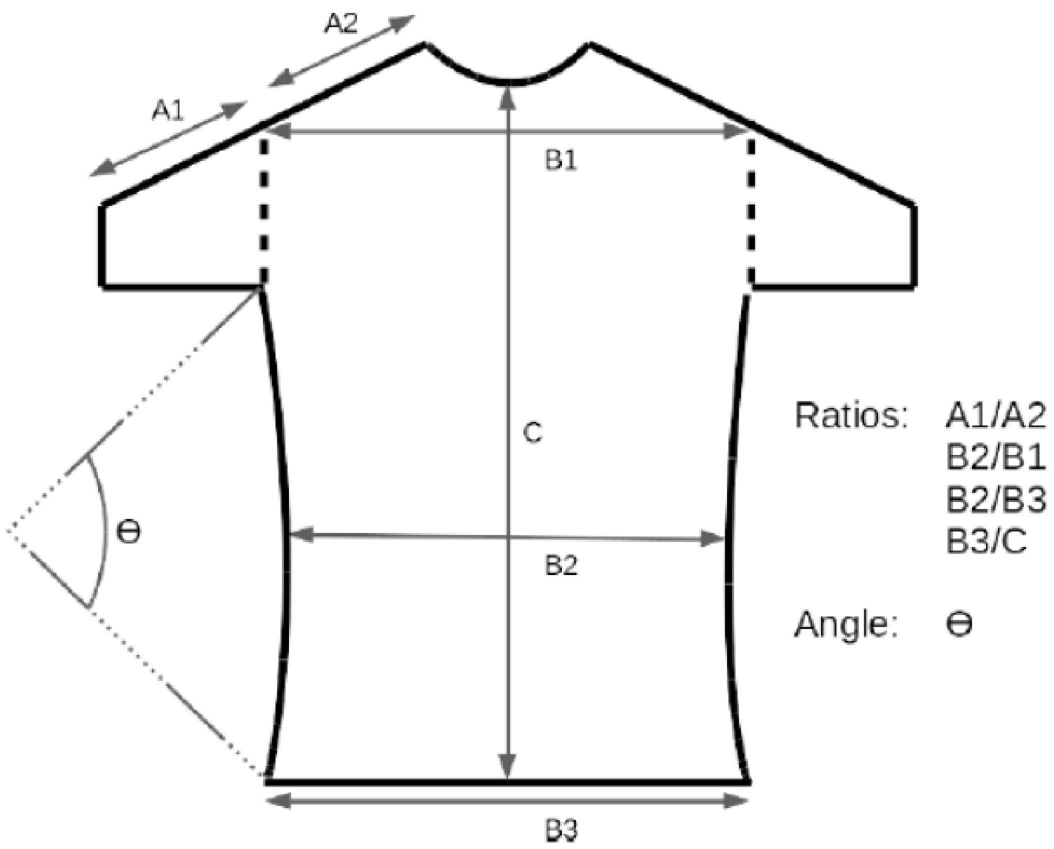

FIG.5

Amit Koshal
IN-PA-2358
Agent for the Applicant

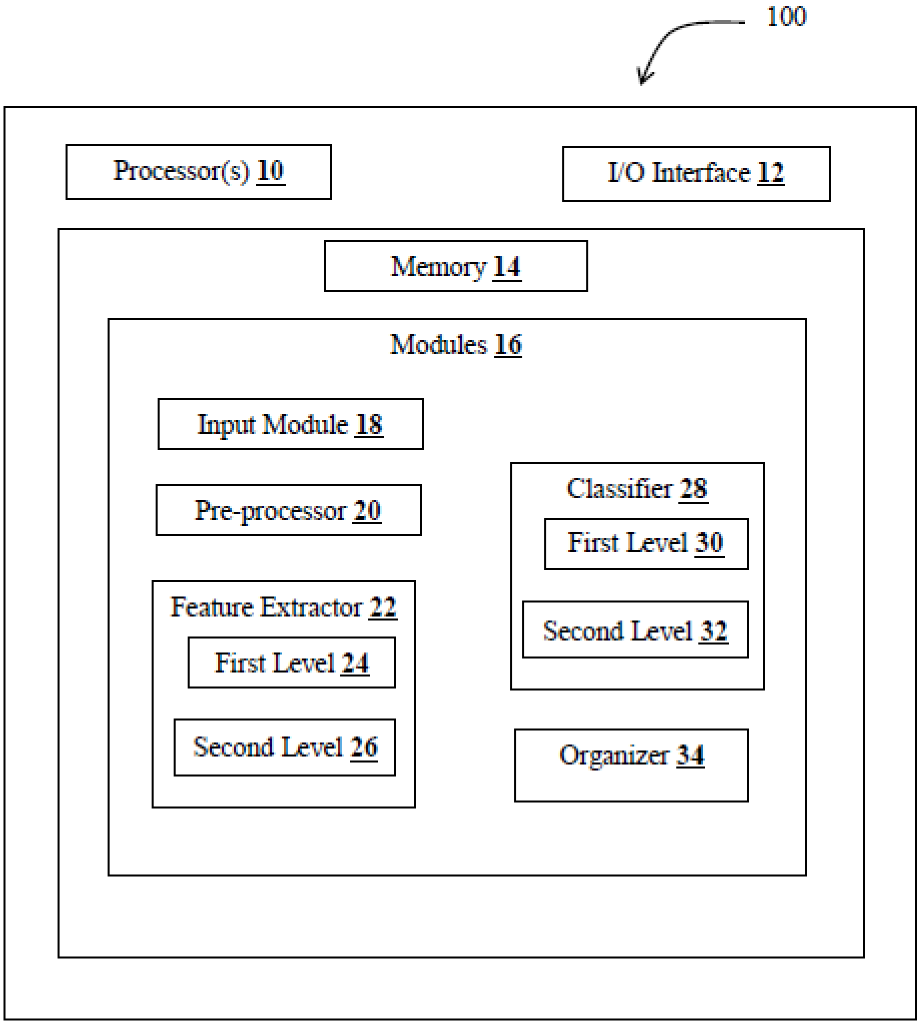

FIG. 1

Amit Koshal
IN-PA-2358
Agent for the Applicant

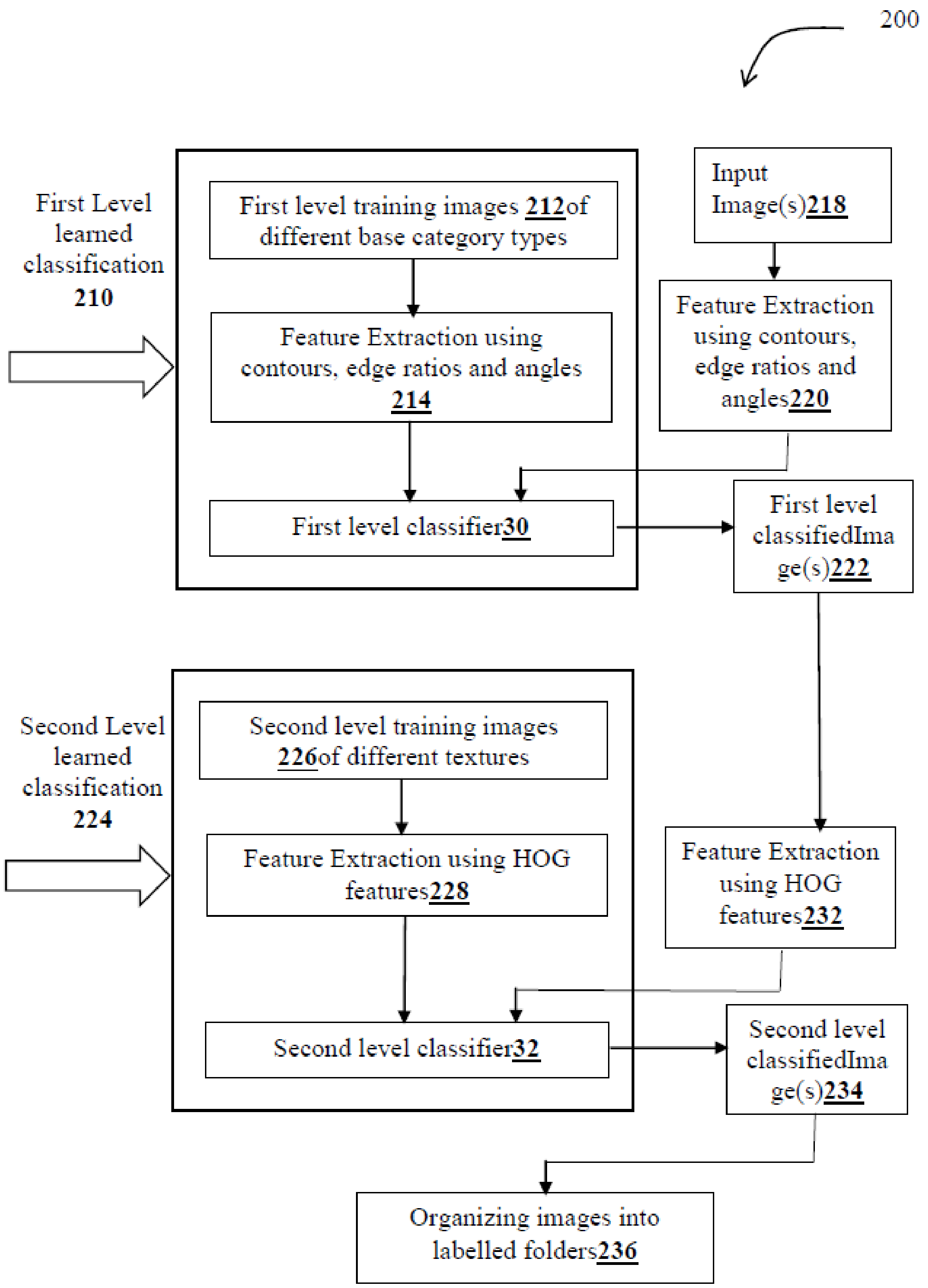

FIG.2

Amit Koshal
IN-PA-2358
Agent for the Applicant

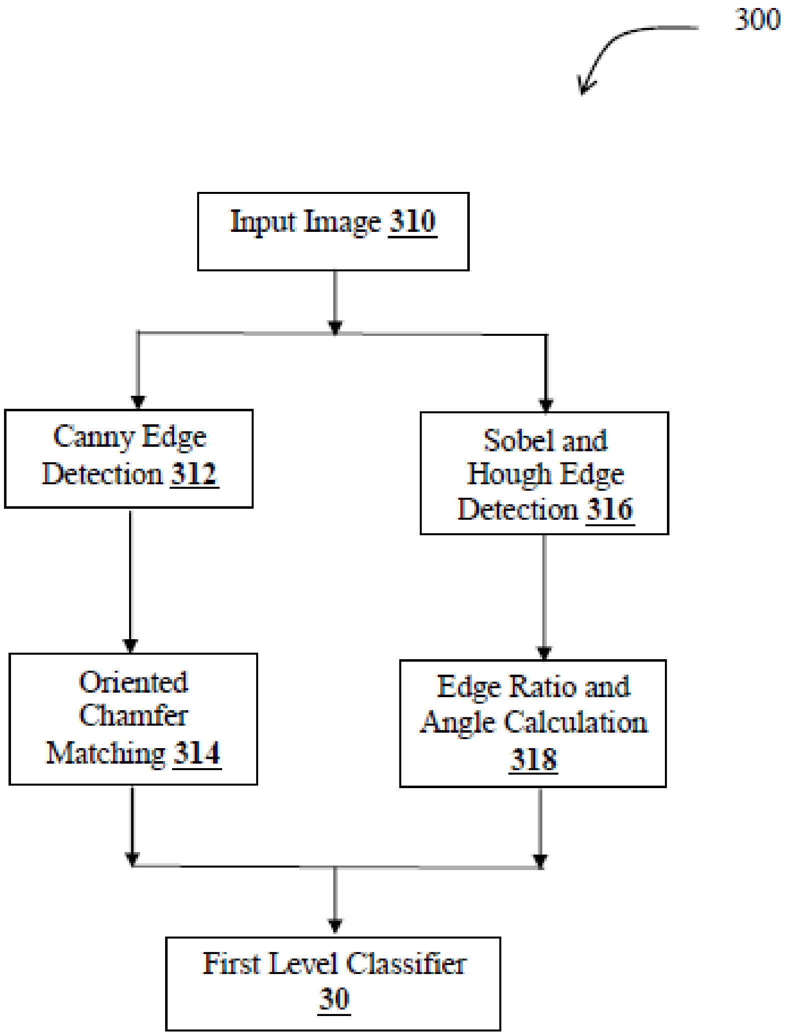

FIG.3

Amit Koshal
IN-PA-2358
Agent for the Applicant

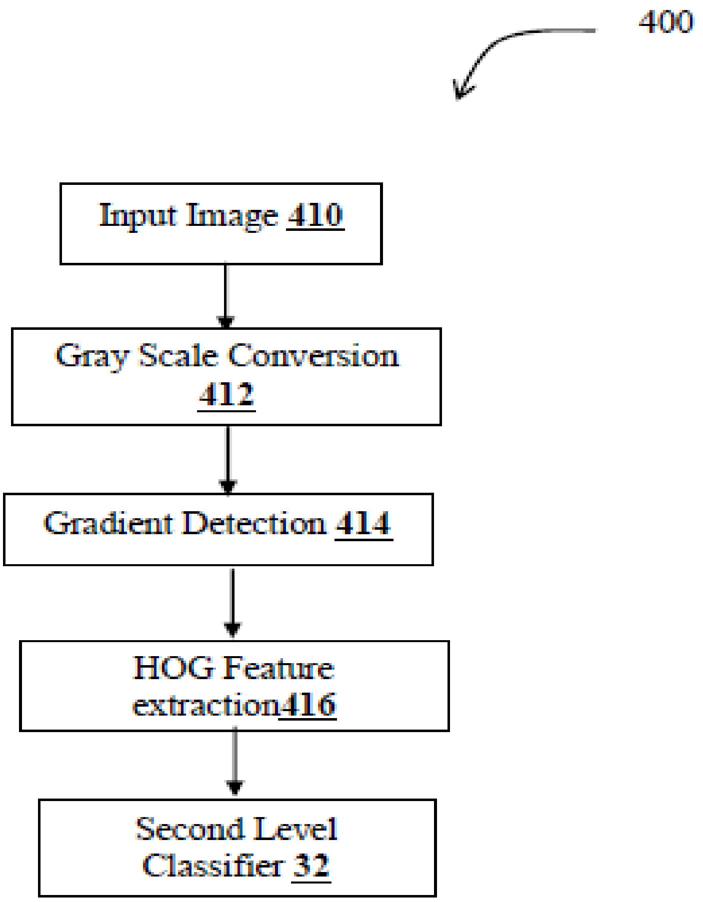

FIG.4

Amit Koshal
IN-PA-2358
Agent for the Applicant

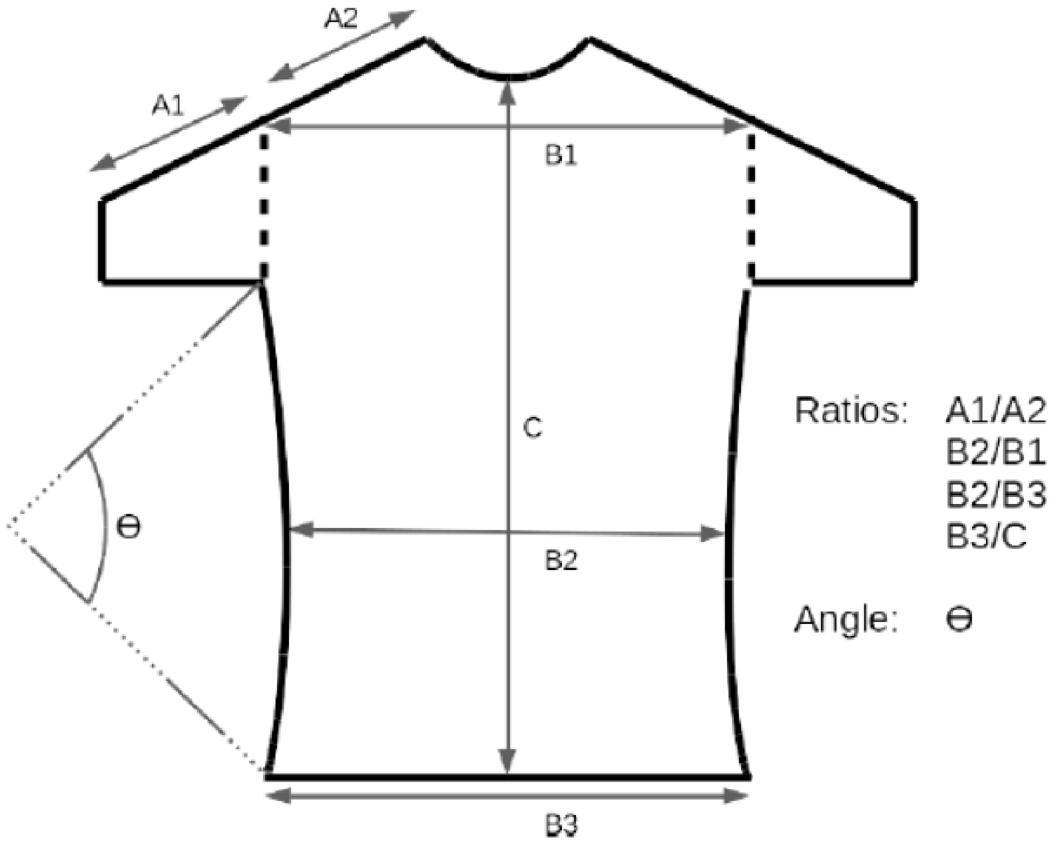

FIG.5

Amit Koshal
IN-PA-2358
Agent for the Applicant