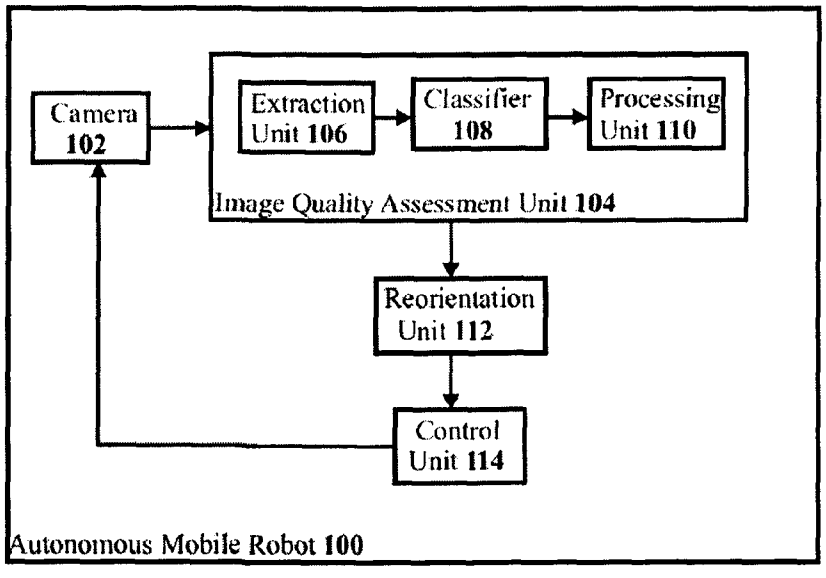
## ABSTRACT

An autonomous robot without stereo vision for capturing high quality and aesthetically good images and a method thereof have been disclosed. The autonomous robot employs an iterative approach to assess the quality of an image based high level features combined with features of 'aesthetic guidelines of professional photography' to automatically re-orient its camera to capture better human as well as non-human images. Thus, the robot captures professional photographs that are in accord with the human professional photography.

**WE CLAIM**

1. An autonomous robot without stereo vision for capturing high quality and aesthetically good images based on 'aesthetic composition guidelines of professional photography', comprising:
    - at least one camera for capturing images;
    - an image quality assessment unit adapted to receive the captured image, said image quality assessment unit comprising:
        - an extraction unit for extracting salient regions from the received image and generating a saliency map to compute high level features of the received image;
        - a classifier adapted to classify the received image as a 'high quality image' and a 'low quality image' based on the computed high level features;
        - a processing unit adapted to apply at least two predetermined features of 'aesthetic composition guidelines of professional photography' on the received image classified as a 'high quality image' and compute an aesthetic deviation value for each of said features;
    - a reorientation unit adapted to receive said computed aesthetic deviation values for said features and determine a change value (Δ□) until said computed aesthetic deviation values are less than predetermined thresholds; and
    - a control unit co-operating with said camera, said control unit adapted to control the orientation of said camera based on the received change value (Δ□) to capture a high quality and aesthetically good image.

2. The autonomous robot as claimed in claim 1, wherein said extraction unit is adapted to generate a saliency map for the received image using visual attention models.

3. The autonomous robot as claimed in claim 1, wherein said extraction unit is further adapted to compute high level features of the received image using thresholding, wherein said high level features are selected from the group consisting of clarity contrast, lighting, simplicity and color harmony.

4. The autonomous robot as claimed in claim 1, wherein said predetermined features of aesthetic composition guidelines of professional photography are selected from the group consisting of rule of thirds and golden ratio rule.

5. The autonomous robot as claimed in claim 1, wherein said reorientation unit is adapted to determine said change value (Δ□) required to reorient said camera in at least one of the four directions by calculating Euclidean distance and Manhattan distance for the extracted salient region of the received image.

6. The autonomous robot as claimed in claim 1, wherein said reorientation unit is adapted to determine said change value (Δ□) required to reorient said camera in at least one of the four directions by computing the depth of the extracted salient region by creating depth maps using depth estimation techniques.

7. A method for facilitating an autonomous robot without stereo vision for capturing high quality and aesthetically good images, said method comprising the following steps:

- capturing an image using at least one camera associated with the autonomous robot;
- extracting salient regions from the captured image and generating a saliency map to compute high level features of the received image;
- classifying the captured image as a 'high quality image' or a 'low quality image' based on the computed high level features;
- applying at least two features from 'aesthetic composition guidelines of professional photography' on the received image classified as a 'high quality image' and computing an aesthetic deviation value for each of said features;
- determining a change value (Δ□) based on said computed aesthetic deviation values until said computed aesthetic deviation values are less than a predetermined threshold; and
- controlling the orientation of said camera based on the received change value (Δ□) to capture a high quality and aesthetically good image.

8. The method as claimed in claim 6, wherein the step of extracting salient regions from the captured image and generating a saliency map to compute high level features includes the following steps:
   - applying visual attention models for generating a saliency map for the captured image; and
   - computing high level features of the received image by thresholding said saliency map using a predetermined threshold value, wherein said high level features are selected from the group consisting of clarity contrast, lighting, simplicity and color harmony.

9. The method as claimed in claim 6, wherein the step of applying at least two features from 'aesthetic composition guidelines of professional photography' on the received image includes the step of applying features including rule of thirds and golden ratio rule to assess the aesthetic quality of the captured image and computing aesthetic deviation values for said rule of thirds and said golden ratio rule features.

10. The method as claimed in claim 6, wherein the step of determining a change value (Δ□) based on said computed aesthetic deviation values includes the following steps:

- calculating the centroid of the subject extracted salient region of the image;
- checking if the calculated centroid of the subject region coincide with four intersection points as per the requirement of rule of thirds feature of 'aesthetic composition guidelines of professional photography';
- calculating the distance required to shift the calculated centriod to the nearest intersection point using the Euclidean distance technique;
- calculating the deviation of a horizontal line of said subject region and the value of ratio as per the requirement of golden ratio rule feature of 'aesthetic composition guidelines of professional photography' using Manhattan distance;
- positioning the camera in the upward direction if said Manhattan distance or the value of ratio is less than a predetermined value else positioning the camera in the downward direction;

- positioning the camera in the left direction if said Euclidean distance is less than a predetermined value else positioning the camera in the right direction; and
- calculating the orientation angle for positioning the camera in the determined direction.

Dated this 22nd day of June, 2012

**MOHAN DEWAN**
**Of R. K. DEWAN& CO.**
**APPLICANTS' PATENT ATTORNEY**

NAME: INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY, HYDERABAD
NO. : 2138/CHE/2011

NO. OF SHEETS:03
SHEET NO. :01

COMPLETE SPECIFICATION

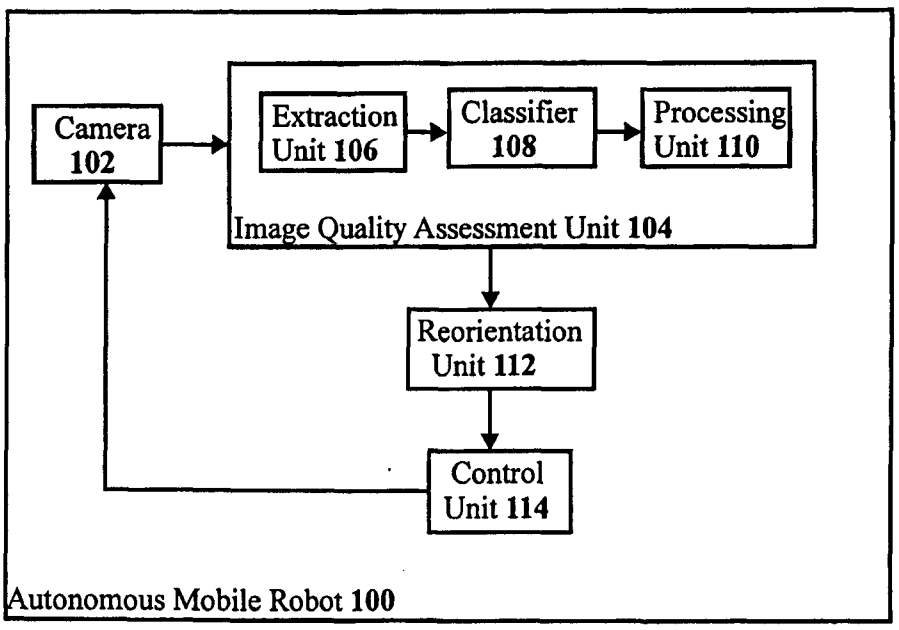



FIGURE 1



(a) (b)

FIGURE 2

MOHAN DEWAN
of R.K. DEWAN & CO.
APPLICANTS' PATENT ATTORNEY

ORIGINAL
2138/CHE/2011

NAME: INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY, HYDERABAD
NO. : 2138/CHE/2011

NO. OF SHEETS:03
SHEET NO. :02

COMPLETE SPECIFICATION

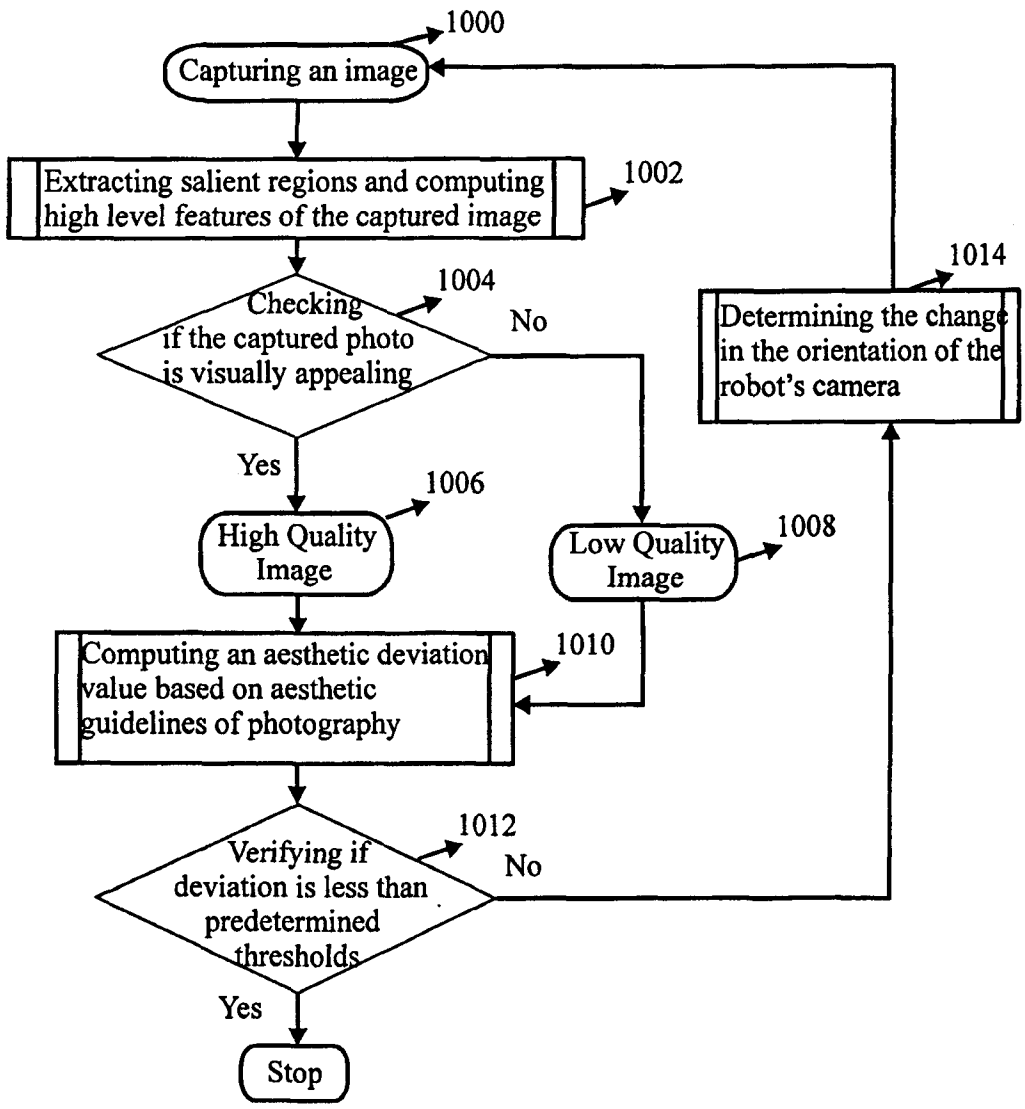



FIGURE 3

MOHAN DEWAN
of R.K. DEWAN & CO.
APPLICANTS' PATENT ATTORNEY

NAME: INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY, HYDERABAD
NO. : 2138/CHE/2011

NO. OF SHEETS:03
SHEET NO. :03

COMPLETE SPECIFICATION

INITIAL PHOTOGRAPH FINAL PHOTOGRAPH

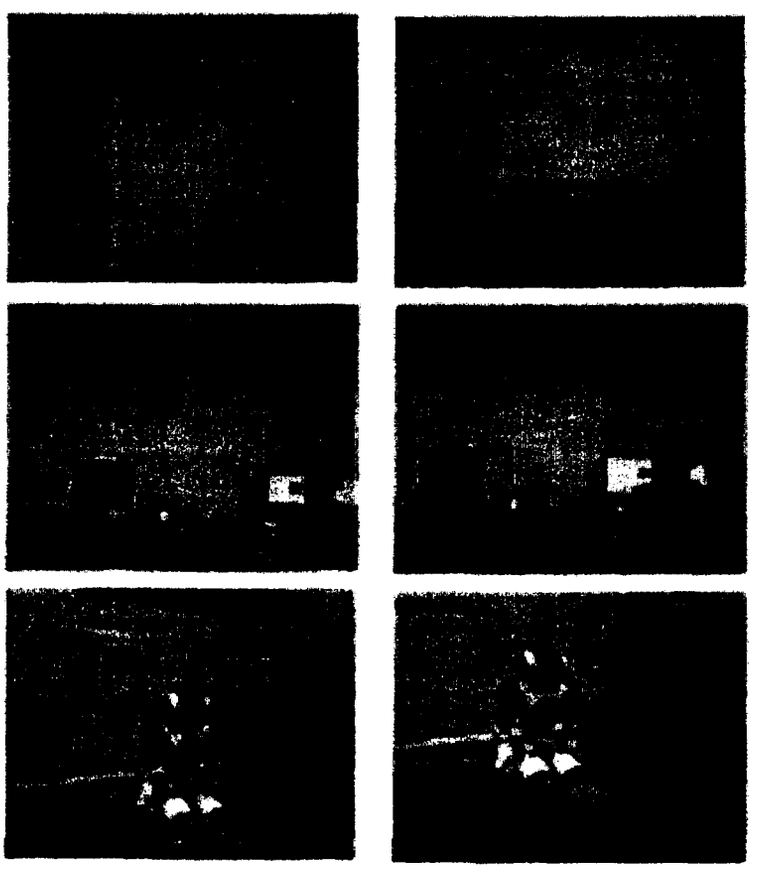


FIGURE 4

MOHAN DEWAN
of R.K. DEWAN & CO.
APPLICANTS' PATENT ATTORNEY

## FIELD OF DISCLOSURE

The present disclosure relates to the field of image processing and robotics.

## DEFINITIONS OF TERMS USED IN THE SPECIFICATION

The term 'salient region' used in this specification relates to a particular area in an image where the camera is focused or the area which significantly stands out from the surrounding visual attributes of an image. The other areas in the image are comparatively blurred and form a part of the background of the image.

The term 'visual attention model' in this specification relates to techniques which aim at simulating the selective attention aspect of the human visual system to extract salient features in an image based on the value of color feature, color intensity feature, orientation feature and motion feature derived from pixels of an image.

The term 'aesthetic guidelines of professional photography' in this specification relates to techniques / rules used by professional photographers for capturing photographs / images.

## BACKGROUND

Autonomous robots are known for their self-sufficiency in performing a desired task. These robots can perform the desired task without human guidance or intervention even in unfamiliar or unstructured environments.

The autonomous robots have been used in various fields to assist human beings and one of the areas of research has been in the field of enabling autonomous robots to take photographs.

The prior art systems for autonomous robots employ computer vision techniques like face recognition, smile recognition or skin recognition to facilitate the robots in taking photographs. However, these prior art systems use computationally intensive techniques and are limited to human photography.

There is therefore felt a need for an autonomous robot for capturing aesthetically good human as well as non-human photographs using computationally effective techniques.

**OBJECTS**

Some of the objects of the present disclosure, which at least one embodiment herein satisfies, are as follows:

It is an object of the present disclosure to provide an autonomous robot which captures aesthetically better photographs.

It is another object of the present disclosure to provide an autonomous robot which eliminates the use of computer vision techniques for capturing photographs.

It is yet another object of the present disclosure to provide a computationally effective technique for capturing aesthetically good photographs using an autonomous robot.

Other objects and advantages of the present disclosure will be more apparent from the following description when read in conjunction with the accompanying figures, which are not intended to limit the scope of the present disclosure.

**SUMMARY**

In accordance with this disclosure there is provided an autonomous robot without stereo vision for capturing high quality and aesthetically good images based on 'aesthetic composition guidelines of professional photography', comprising:

- at least one camera for capturing images;
- an image quality assessment unit adapted to receive the captured image, the image quality assessment unit comprising:
    - an extraction unit for extracting salient regions from the received image and generating a saliency map to compute high level features of the received image;
    - a classifier adapted to classify the received image as a 'high quality image' and a 'low quality image' based on the computed high level features;
    - a processing unit adapted to apply at least two predetermined features of 'aesthetic composition guidelines of professional photography' on the received image classified as a 'high quality image' and compute an aesthetic deviation value for each of the features;
- a reorientation unit adapted to receive the computed aesthetic deviation values for the features and determine a change value ($\Delta\square$) until the computed aesthetic deviation values are less than predetermined thresholds; and

- a control unit co-operating with the camera, the control unit adapted to control the orientation of the camera based on the received change value (Δ□) to capture a high quality and aesthetically good image.

Typically, the extraction unit is adapted to generate a saliency map for the received image using visual attention models.

Preferably, the extraction unit is further adapted to compute high level features of the received image using thresholding, wherein the high level features are selected from the group consisting of clarity contrast, lighting, simplicity and color harmony.

Further, the predetermined features of aesthetic composition guidelines of professional photography are selected from the group consisting of rule of thirds and golden ratio rule.

Still further, the reorientation unit is adapted to determine the change value (Δ□) required to reorient the camera in at least one of the four directions by calculating Euclidean distance and Manhattan distance for the extracted salient region of the received image. Alternatively, the reorientation unit is adapted to determine the change value (Δ□) required to reorient the camera in at least one of the four directions by computing the depth of the extracted salient region by creating depth maps using depth estimation techniques.

In accordance with this disclosure, there is a provided a method for facilitating an autonomous robot without stereo vision for capturing high quality and aesthetically good images, the method comprising the following steps:

- capturing an image using at least one camera associated with the autonomous robot;

- extracting salient regions from the captured image and generating a saliency map to compute high level features of the received image;
- classifying the captured image as a 'high quality image' or a 'low quality image' based on the computed high level features;
- applying at least two features from 'aesthetic composition guidelines of professional photography' on the received image classified as a 'high quality image' and computing an aesthetic deviation value for each of the features;
- determining a change value (Δ□) based on the computed aesthetic deviation values until the computed aesthetic deviation values are less than a predetermined threshold; and
- controlling the orientation of the camera based on the received change value (Δ□) to capture a high quality and aesthetically good image.

Typically, the step of extracting salient regions from the captured image and generating a saliency map to compute high level features includes the following steps:

- applying visual attention models for generating a saliency map for the captured image; and
- computing high level features of the received image by thresholding the saliency map using a predetermined threshold value, wherein the high level features are selected from the group consisting of clarity contrast, lighting, simplicity and color harmony.

Preferably, the step of applying at least two features from 'aesthetic composition guidelines of professional photography' on the received image includes the step of applying features including rule of thirds and golden ratio rule to assess the aesthetic quality of the captured image and computing aesthetic deviation values for the rule of thirds and the golden ratio rule features.

In addition, the step of determining a change value (Δ□) based on the computed aesthetic deviation values includes the following steps:

- calculating the centroid of the subject extracted salient region of the image;
- checking if the calculated centroid of the subject region coincide with four intersection points as per the requirement of rule of thirds feature of 'aesthetic composition guidelines of professional photography';
- calculating the distance required to shift the calculated centriod to the nearest intersection point using the Euclidean distance technique;
- calculating the deviation of a horizontal line of the subject region and the value of ratio as per the requirement of golden ratio rule feature of 'aesthetic composition guidelines of professional photography' using Manhattan distance;
- positioning the camera in the upward direction if the Manhattan distance or the value of ratio is less than a predetermined value else positioning the camera in the downward direction;
- positioning the camera in the left direction if the Euclidean distance is less than a predetermined value else positioning the camera in the right direction; and
- calculating the orientation angle for positioning the camera in the determined direction.

## BRIEF DESCRIPTION OF THE ACCOMPANYING DRAWINGS

The autonomous robot for capturing high quality and aesthetically good images without stereo vision will now be described with reference to the non-limiting, accompanying drawings, in which:

**FIGURE 1** illustrates a block diagram of the autonomous robot for capturing high quality and aesthetically good images without using stereo vision in accordance with the present disclosure;

**FIGURE 2a** shows a sample high quality image and **FIGURE 2b** shows the extracted salient region for the image in accordance with this disclosure;

**FIGURE 3** illustrates a flowchart showing the steps involved in facilitating an autonomous robot for capturing aesthetically good photographs in accordance with the present disclosure; and

**FIGURE 4** shows sample initial and final photographs captured by the proposed disclosure using an autonomous robot in accordance with the present disclosure.

## DETAILED DESCRIPTION OF THE ACCOMPANYING DRAWINGS

A preferred embodiment will now be described in detail with reference to the accompanying drawings. The preferred embodiment does not limit the scope and ambit of the disclosure. The description provided is purely by way of example and illustration.

The embodiments herein and the various features and advantageous details thereof are explained with reference to the non-limiting embodiments in the following description. Descriptions of well-known components and processing techniques are

omitted so as to not unnecessarily obscure the embodiments herein. 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.

The following description of the specific embodiments will so fully reveal the general nature of the embodiments herein that others can, by applying current knowledge, readily modify and/or adapt for various applications such specific embodiments without departing from the generic concept, and, therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments. It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of limitation. Therefore, while the embodiments herein have been described in terms of preferred embodiments, those skilled in the art will recognize that the embodiments herein can be practiced with modification within the spirit and scope of the embodiments as described herein.

The prior art systems based on autonomous robots used computer vision techniques for capturing photographs, these techniques required complex computations and were confined to capturing human photographs. These shortcomings of the prior art led the present disclosure to envisage an autonomous robot to facilitate capturing of aesthetically good photographs which do not rely on the subject of the image being captured. Thus, the envisaged autonomous robot can capture both human and non-human based photographs/ images that are coherent with human perception.

In accordance with one aspect of the present disclosure, the proposed mobile autonomous robot captures aesthetically better photographs without having stereo vision based capabilities. The robot combines high level features of the image with aesthetic guidelines of professional photography to capture professional photographs which are in accord with human professional photography.

Moreover, the robot uses an iterative photo capture mechanism which assesses the captured photograph's quality by detecting aesthetic errors to reposition its camera to capture good quality photographs.

Thus, the disclosure enhances the robot's Focus of Attention (FOA) to capture high quality human and non-human photographs using computationally effective and non-computer vision based techniques.

Referring to the accompanying drawings, **FIGURE 1** illustrates a schematic of the autonomous robot **100** in accordance with this disclosure. The autonomous robot **100** includes at least one camera **102** for capturing aesthetic guideline driven photographs / images based on its current camera orientation. The visual quality of the image is assessed by an image quality assessment unit **104** of the robot **100**. Based on the assessed visual quality a desired change in the orientation of the robot's camera **102** is determined by a reorientation unit **112** and accordingly the robot's camera is repositioned by a control unit **114** to capture better images.

In accordance with this disclosure, the image quality assessment unit **104** includes an extraction unit **106** which receives a captured image from the robot's camera **102**. The extraction unit **106** extracts salient regions in the captured image using visual attention models and generates a saliency map. This is contrary to the technique followed in the prior art of extracting blurred regions and subtracting

them from the original image to retrieve the salient regions in an image. In accordance with this disclosure, the basic principle behind extracting salience is the detection of locations in the image whose visual attributes significantly differ or stand out from the surrounding attributes. **FIGURE 2a** shows an exemplary image and its corresponding salient region extracted by the extraction unit **106** in **FIGURE 2b**.

The saliency map is further used by the extraction unit **106** to compute high level features of the received image. The high level features extracted by extraction unit **106** include clarity contrast feature ($f_c$), lighting feature ($f_l$), simplicity feature ($f_s$) and color harmony feature ($f_h$). These high level features are provided to a classifier **108**. In accordance with this disclosure, the classifier **108** is a basic two class SVM classifier which learns the aforementioned high level features and runs them on the received image to judge its visual appeal and classify it as a high quality image or a low quality image. The classifier **108** does not require any reference image quality techniques to determine the visual quality of the photograph / image in accordance with this disclosure.

The classified 'high quality image' or 'low quality image' is given to a processing unit **110**. The processing unit **110** based on predetermined features of aesthetic guidelines of photography calculates a set of aesthetic deviation values for the classified image. The present disclosure proposes the use of two well known composition guidelines of professional photography namely the rule of thirds and the golden ratio rule to assess the quality of the captured image.

The rule of thirds feature of professional photography suggests that an image must be imagined as divided into nine equal parts by two equally-spaced horizontal lines and two equally spaced vertical lines, and that important compositional elements

must be placed along these lines or their intersections (i.e. intersection points). As per the rule of thirds feature, aligning a subject with these points creates more tension, energy and interest in the composition than simply centering only the subject. The golden ratio rule feature of professional photography suggests that the ratio between the areas of the rectangles formed because of the horizon line must be equal to the golden mean, which is having a value of 1.618, to be more pleasing to the eye.

In accordance with this disclosure, the processing unit **110** applies the two features - rule of thirds and golden ratio rule on the received image and calculates the aesthetic deviation value for them.

The aesthetic deviation value for rule of thirds feature ($f_{th}$) is calculated as the minimum deviation observed by the centroid (Cx, Cy) of the extracted subject region using the below formula:

$$f_{th} = \min_{i=1,2,3,4} \{\sqrt{(C_x - P_{ix})^2/X^2 + (C_y - P_{iy})^2/Y^2}\}$$

where (Pix, Piy) are the four intersection points of the image for values of i = 1, 2, 3, 4 and X and Y are width and height of the image respectively. The value $f_{th}$ has a bounded range between [0, 0.47], as the maximum deviation from rules of thirds occur when the centroid of subject region coincides with any of the corners of the image.

The aesthetic deviation value for golden ratio feature (fgr), is calculated by computing the ratio (r) of areas of the rectangles formed by the horizon line of the image which is generated using the vanishing point detector using the below formula:

$$f_{gr} = |\max\{r, 1/r\} - 1.618|$$

In accordance with this disclosure, large values of $f_{th}$ or $f_{gr}$ indicate high aesthetic deviation of the photograph / image. Further, the composition guidelines are followed if the $f_{th}$, $f_{gr}$ feature values are less than certain thresholds. These thresholds are determined by taking the average of the corresponding feature values computed on a dataset of good professional photographs. In accordance with this disclosure, the thresholds set for $f_{th}$ and $f_{gr}$ are typically, 0.25 and 0.30 respectively.

The computed aesthetic deviation values and the predetermined thresholds are given to the reorientation unit **112**. If the aesthetic deviation values are below the predetermined aesthetic parameter thresholds, typically 0.25 for $f_{th}$ and 0.30 for $f_{gr}$ then the captured photograph / image is aesthetically good and does not require the robot to reposition its camera otherwise the reorientation unit **112** calculates a change value (Δ□).

To satisfy the rule of thirds, the reorientation unit **112** uses the centroid (CRx, CRy) of the subject region and checks if the centroid coincides with any of the four intersecting points (Pi). The point nearest to the centroid region is chosen by the reorientation unit **112** for calculating the Euclidean distance as seen below:

$$distance = \min_{i=1,2,3,4} \{\sqrt{(C_x - P_{ix})^2 + (C_y - P_{iy})^2}\}$$

The Euclidean distance represents the distance by which the centroid of the subject salient region has to be shifted to the desired location. Accordingly, the orientation of the robot needs to be changed with a certain angle ($\Delta\theta = (\Delta\theta x, \Delta\theta y)$) along the axes of the captured image to coincide with the nearest of the four points of the thirds rule.

For images following the golden ratio rule, the reorientation unit **112** calculates the deviation of the horizon line using the Manhattan distance of corresponding points on the deviated and the aesthetical horizon lines.

The reorientation unit **112** then compares the calculated Euclidean distance ($C_y$ - $P_y$) and the Manhattan distance ($C_x$-$P_x$) with predetermined values to determine the direction of orientation of the camera in top, bottom, left and right directions and set the at least one determined direction as the value of change ($\Delta\theta$).

Thus, the reorientation unit **112** performs the following steps for the camera reorientation where ($C_x$; $C_y$), ($P_x$; $P_y$) are the centroid of subject region and the nearest point from rule of thirds respectively and r is the ratio of areas of upper rectangle to the lower rectangle formed by the horizon line drives the robot re-orientation.

- Setting the value of □ = (Angle of the 'view range' of the robot camera)/2
- Checking if $f_{th} > 0{:}25$ or $f_{gr} > 0{:}30$
  - Checking if $(C_x - P_x) < 0$ or $r < 0{:}618$ or ($r > 1$and $r < 1{:}618$)
    - then orienting the camera ↑ (top)
    - else orienting the camera ↓ (bottom)
  - if $(C_y - P_y) < 0$
    - then orienting the camera ← (left)
    - else orienting the camera → (right)
- Reorient the robot camera □ = □/2 along the required direction;
- recapturing a new photograph;
- calculating $f_{th}$, $f_{gr}$ and reiterating the steps until an image with less aesthetic deviation is captured.

A naive approach to reorient the robot can be done by changing the orientation of the robot camera in integral multiples of small angle (□□), say $1^0$. The problem with this approach is the error in the movement of the robot camera gets compounded, which may sometimes result in much more deviated photographs. Also the number of intermediate images captured increases linearly with the deviation. To reduce the compounded error and reorient the robot in reduced time, the present disclosure follows an approach which is logarithmically converging to capture the required photograph. For a given angle view range of the robot camera, the number of photographs taken is bounded by the below formula:

$$\lceil \log_2(angle_view_range) \rceil - 1$$

The reorientation unit **112** forwards the determined change ($\Delta\theta$) to the control unit **114** which further instructs the robot to reposition its camera **102** and capture a new image.

In accordance with this disclosure, the change in robot reorientation can also be determined by computing the depth of the extracted salient region using depth estimation algorithms from computer vision.

In accordance with this disclosure, there is provided a method for facilitating an autonomous robot without stereo vision for capturing high quality and aesthetically good images as seen in **FIGURE 3**, the method comprising the following steps:

- capturing an image using at least one camera associated with the autonomous robot **1000**;
- extracting salient regions from the captured image and generating a saliency map to compute high level features of the received image **1002**;
- classifying the captured image as a 'high quality image' **1006** or a 'low quality image' **1008** based on the computed high level features **1004** ;

- applying at least two features from 'aesthetic composition guidelines of professional photography' on the received image classified as a 'high quality image' and computing an aesthetic deviation value for each of the features **1010**;
- determining a change value (Δ□) based on the computed aesthetic deviation values until the computed aesthetic deviation values are less than a predetermined threshold **1012**; and
- controlling the orientation of the camera based on the received change value (Δ□) to capture a high quality and aesthetically good image **1014**.

Typically, the step of extracting salient regions from the captured image and generating a saliency map to compute high level features includes the following steps:

- applying visual attention models for generating a saliency map for the captured image; and
- computing high level features of the received image by thresholding the saliency map using a predetermined threshold value, wherein the high level features are selected from the group consisting of clarity contrast, lighting, simplicity and color harmony.

Preferably, the step of applying at least two features from 'aesthetic composition guidelines of professional photography' on the received image includes the step of applying features including rule of thirds and golden ratio rule to assess the aesthetic quality of the captured image and computing aesthetic deviation values for the rule of thirds and the golden ratio rule features.

In addition, the step of determining a change value (Δ□) based on the computed aesthetic deviation values includes the following steps:

- calculating the centroid of the subject extracted salient region of the image;
- checking if the calculated centroid of the subject region coincide with four intersection points as per the requirement of rule of thirds feature of 'aesthetic composition guidelines of professional photography';
- calculating the distance required to shift the calculated centriod to the nearest intersection point using the Euclidean distance technique;
- calculating the deviation of a horizontal line of the subject region and the value of ratio as per the requirement of golden ratio rule feature of 'aesthetic composition guidelines of professional photography' using Manhattan distance;
- positioning the camera in the upward direction if the Manhattan distance or the value of ratio is less than a predetermined value else positioning the camera in the downward direction;
- positioning the camera in the left direction if the Euclidean distance is less than a predetermined value else positioning the camera in the right direction; and
- calculating the orientation angle for positioning the camera in the determined direction.

**EXPERIMENTAL DETAILS**

The quality assessment approach of the present disclosure was tested using experiments which were conducted on NAO robot with no stereo vision. The NAO can rotate its head in all four directions up, down, left and right. And the angle view range of the NAO's camera is $34^0$. The NAO was trained on a large and diversified dataset containing 6000 good and 6000 bad images. The robot was able to achieve a classification accuracy of about 79% on this dataset. The robot was

able to assess the aesthetic error (deviation) in the captured image and estimate the change required in orientation of the camera to retake an aesthetically better photograph. **FIGURE 4** shows the initial photographs which were taken by a NAO robot and the corresponding final photographs which were captured after re-orientation of the robot's camera using the present disclosure. The results demonstrate that the proposed robot can be used to capture professional photographs which are in accord with the human professional photography.

## TECHNICAL ADVANTAGES

The technical advancements of the present disclosure include realization of an autonomous robot for capturing high quality and aesthetically good images without stereo vision that are coherent with human perception.

The robots use saliency based no-reference image quality assessment techniques to verify the quality of the images/photographs captured. Using visual salience for image quality assessment reduces the computational complexity of the image quality assessment process followed by the robots and provides a faster and efficient method for capturing aesthetically better photographs.

In addition, the need for using complex computer vision based techniques is eliminated and reference image based quality assessment techniques is used to provide a relatively inexpensive and rapid mechanism to detect aesthetic errors and accordingly reorient the robot's camera to capture good photographs.

Due to its efficiency and speed in capturing aesthetically good photographs the proposed disclosure can in applied in following applications including:

- in surveillance systems where the camera is continuously rotating to capture good photographs;

- in web image search applications to rank images;
- in professional photography using robot photographers;
- for navigation of the robot;
- in upcoming domains like robot-journalism and paparazzi; and
- in locations inaccessible by humans like difficult terrains.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

The use of the expression "at least" or "at least one" suggests the use of one or more elements or ingredients or quantities, as the use may be in the embodiment of the invention to achieve one or more of the desired objects or results.

Any discussion of documents, acts, materials, devices, articles or the like that has been included in this specification is solely for the purpose of providing a context for the invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the invention as it existed anywhere before the priority date of this application.

Wherever a range of values is specified, a value up to 10% below and above the lowest and highest numerical value respectively, of the specified range, is included in the scope of the disclosure.

The numerical values mentioned for the various physical parameters, dimensions or quantities are only approximations and it is envisaged that the values higher/lower than the numerical values assigned to the parameters, dimensions or quantities fall within the scope of the disclosure, unless there is a statement in the specification specific to the contrary.

While considerable emphasis has been placed herein on the particular features of this invention, it will be appreciated that various modifications can be made, and that many changes can be made in the preferred embodiment without departing from the principles of the invention. These and other modifications in the nature of the invention or the preferred embodiments will be apparent to those skilled in the art from the disclosure herein, whereby it is to be distinctly understood that the foregoing descriptive matter is to be interpreted merely as illustrative of the invention and not as a limitation.