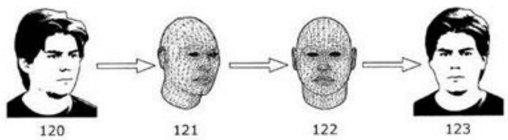
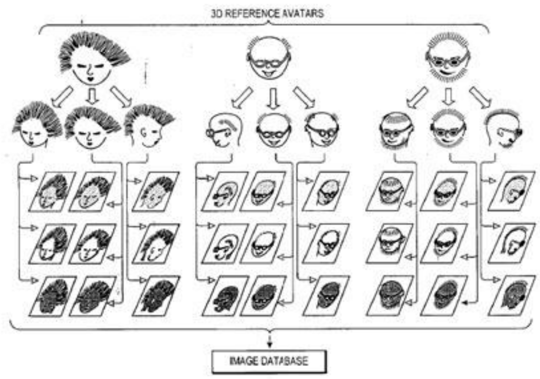
Animetrics Patents

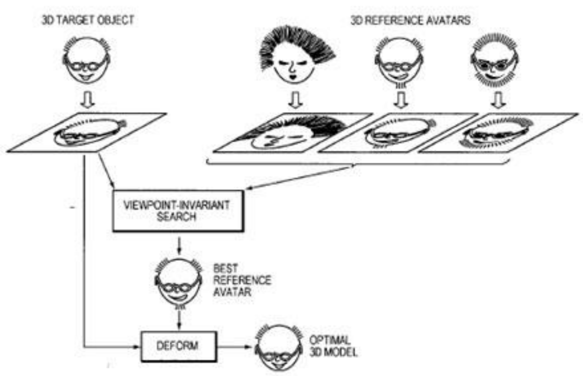
Facial Recognition System and Method
US 7,643,671 B2 Jan.5, 2010



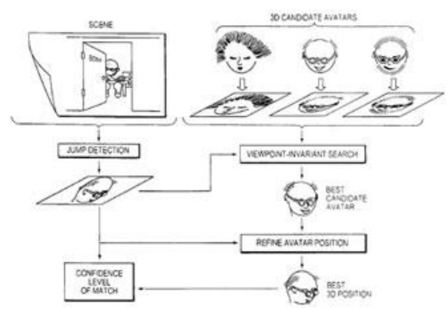
Generation of Image Database for Multifeatured Objects
US 7,643,683 B2 Jan.5, 2010



Viewpoint-Invariant Image Matching and Generation of 3D Models from 2D Imagery
US 7,643,685 B2 Jan.5, 2010



Viewpoint-Invariant Detection And ID of 3D Object from 2D Imagery
US 7,853,085 B2 Dec.14,2010



FACIAL RECOGNITION SYSTEM AND METHOD –

US 7,643,671 B2

The present invention substantially overcomes the deficiencies of the prior art by providing a facial recognition system which processes images to correct for lighting and pose prior to comparison. The images are corrected for lighting and pose by using shape information. The system processes a two dimensional image of a face to create a three dimensional image of the face. The three dimensional image is manipulated to change the pose and lighting characteristics. Finally, the modified three dimensional image is converted back to a two dimensional image prior to processing for recognition. The three dimensional image is manipulated to be facing forward and with a diffuse light from the front.

A plurality of two dimensional images of faces of individuals are processed to create three dimensional images. The three dimensional images are manipulated to change the pose and lighting characteristics to achieve uniformity among the plurality of images. The three dimensional images are stored in a database for later comparison with another image. The three dimensional images are converted into two dimensional images before being stored in a database. The three dimensional images are manipulated to be facing forward and with a diffuse light from the front. A two dimensional image is compared with at least one of the images in the database to determine whether a match exists. A two dimensional image is compared to a set of images in the database to determine a relative value of a match.

A facial recognition system compares a newly acquired image of a face to images of faces in a database to determine a match. The newly acquired image includes one or more two dimensional images. The system processes the one or more two dimensional images to create a three dimensional image of the face. The three dimensional image is manipulated to change the pose and lighting characteristics. According to an aspect of the invention, the three dimensional image is manipulated to be facing forward and with a diffuse light from the front. The three dimensional image is processed to create a second two dimensional image. The second two dimensional image is compared with images in the database to determine whether a match exists.

An iterative process is used to create the three dimensional image from the original two dimensional image. An initial shape is used with data from the two dimensional image to create a three dimensional shape. The three dimensional shape is iteratively adjusted to match the original image. According to an aspect of the invention, at each iteration, a two dimensional image is rendered from the three dimensional image. The rendered two dimensional image is compared to the original image. The three dimensional shape is adjusted based upon the differences between the images. The three dimensional shape is determined based upon estimates of lighting in the original image. The three dimensional shape is determined based upon estimates of pose in the original image.

**Generation of Image Database for Multifeatured Objects -
US 7,643,683 B2**

The present invention provides a method and system for creating a database of images of multifeatured objects. The image database may be used for the purpose of training an identification system or for use directly in identification. The technique uses one or more 3D representations of the multifeatured object(s). In the case where a single individual object is to be identified, a single 3 D representation corresponding to the object is used to generate the database. If more than one individual object is to be identified, the database is generated by combining images from 3D representations of each individual, if they are available. An image database may also be created when a 3D representation of the particular object to be identified is not available. One or more 3D representations of generically similar objects may be used instead. An image database for training or identifying any member of a generic type may be created by combining images generated from a large set of representations of objects of the generic type.

The images may be generated from the 3D representation by producing 2D projections from a range of viewpoints and under a range of lighting conditions. In addition, small or large deformations of each 3D representation corresponding to anticipated internal movements of the target object can be used to generate projections. The database may be indexed and classified using known methods. For example, indexing may be performed according to pose (e.g., position in three dimensions, rotations in three dimensions, and deformation). The classification of images may be hierarchical, and probes may be generated using Bayesian classifiers, neural nets, or other techniques. In a first aspect, therefore, the invention comprises a method of creating a database of images of multifeatured objects of a generic type. In accordance with the method, at least one reference 3D representation of a multi featured object of the generic type is provided. The reference 3D representation(s) may include a representation that corresponds to at least one of the target multi featured objects. The reference 3D representation(s) are used to generate a plurality of images. The image database is then populated with the generated images. In one embodiment, the images are generated from a range of small or large deformations of each 3D representation which accommodate movements of part of the multifeatured object. In another embodiment, the images are generated by creating projections of the reference 3D representation(s) and their deformed variations from a range of viewpoints. In yet another embodiment, a range of images corresponding to varying lighting conditions is generated for each projection. In yet another embodiment, a range of images corresponding to varying textured color conditions of the reference 3D representation is generated for each projection. These embodiments may be variously combined.

In a second aspect, the invention comprises a system for creating a database of images of multi featured objects of a generic type. The system comprises a generator database comprising at least one reference 3D representation of a multifeatured object of the generic type, an image generator for generating a plurality of images from the reference 3D representation(s) in the generator database, and an image database for storing the generated images. The generator database may include a representation that corresponds to at least one of the target objects. In one embodiment, the image generator generates a range of small or large deformations of the 3D representation to accommodate movements of parts of the multifeatured object. In another embodiment, the image generator generates 2D projections of the reference 3D representation(s) and their deformed variants from a range of viewpoints. In yet another embodiment, the image generator generates a range of images corresponding to varying lighting conditions for each viewpoint. In yet another embodiment, the image generator generates a range of images corresponding to varying textured color conditions for each viewpoint. These embodiments may be variously combined. In a third aspect, the above described methods and systems are used for the case when the 3D multifeatured object is a face, and the 3D representations are "avatars," i.e., electronic 3D graphical or pictorial representations.

VIEWPOINT-INVARIANT IMAGE MATCHING AND GENERATION OF THREE-DIMENSIONAL MODELS FROM TWO-DIMENSIONAL IMAGERY –

US 7,643,685 B2

The present invention provides an automated method and system for generating an optimal 3D model of a target multifeatured object when only partial source data describing the object is available. The partial source data often consists of one or more 2D projections of the target object or an obscuration of a single projection, but may also include 3D data, such as from a 3D camera or scanner. The invention uses a set of reference 3D representations that span, to the extent practicable, the variations of the class of objects to which the target object belongs. The invention may automatically identify feature items common to the source data and to the reference representations, and establish correspondences between them. For example, if the target object is a face, the system may identify points at the extremities of the eyes and mouth, or the nose profile, and establish correspondences between such features in the source data and in the reference representations. Manual identification and matching of feature items can also be incorporated if desired. Next, all possible positions (i.e., orientations and translations) for each 3D reference representation are searched to identify the position and reference representation combination whose projection most closely matches the source data. The closeness of match is determined by a measure such as the minimum mean squared error (MMSE) between the feature items in the projection of the 3D representation, and the corresponding feature items in the source projection. A comparison is performed in 3D between the estimated projected positions of the feature items from the 2D source projection and the corresponding feature items of the 3D representation. The closest-fitting 3D reference representation may then be deformed to optimize the correspondence with the source projection. Each point in the mesh which defines the geometry of the 3D representation is free to move during the deformation. The search for the best-fitting position (i.e., orientation and translation) is repeated using the deformed 3D representation, and the deformation and search may be repeated iteratively until convergence occurs or terminated at any time.

Thus the geometry of the 3D model is tailored to the target object in two ways. First, when more than one reference representation is available, the selection of the best-fitting reference representation from a set of references enables the optimal coarse-grain choice to be made. Second, deformation enables fine scale tuning in which errors introduced by inaccurate choice of viewpoint are progressively reduced by iteration. The invention requires no information about the viewpoint from which the 2D source projection was captured, because a search is performed over all possible viewpoints, and the viewpoint is taken to be that which corresponds to the closest fit between the projected 3D representation and the 2D source data.
In a first aspect, the invention comprises a method of comparing at least one

source 2D projection of a source multifeatured object to a reference library of3D reference objects. In accordance with the method, a plurality of reference 3D representations of generically similar multifeatured objects is provided, and a viewpoint-invariant search of the reference 3D representations is performed to locate the reference 3D representation having a 2D projection most resembling the source projection(s). In some embodiments, resemblance is determined by a degree of alignment between feature items in the 3D representation and corresponding feature items in the source 2D projection(s). Each reference 3D representation may be searched over a range of possible 2D projections of the 3D representation without actually generating any projections. The search over a range of possible 2D projections may comprise computing a rigid motion of the reference 3D representation optimally consistent with a viewpoint of the source multi featured object in at least one of the 2D projections. The rigid motions may comprise pitch, roll, yaw, and translation in three dimensions. Automatic camera calibration may be performed by estimation of camera parameters, such as aspect ratio and field of view, from image landmarks.

In some embodiments, the optimum rigid motion may be determined by estimating a conditional mean pose or geometric registration as it relates to feature items comprising points, curves, surfaces, and sub volumes in a 3D coordinate space associated with the reference 3D representation such that the feature items are projectionally consistent with feature items in source 2D projection(s). MMSE estimates between the conditional mean estimate of the projected feature items and corresponding feature items of the reference 3D representation are generated. The rigid motion may be constrained by known 3D position information associated with the source 2D projection(s).

In some embodiments, the feature items may include curves as well as points which are extracted from the source projection using dynamic programming. Further, areas as well as surfaces and or subvolumes may be used as features generated via iso contouring (such as via the Marching Cubes algorithm) or automated segmentation algorithms. The feature items used in the matching process may be found automatically by using correspondences between the 2D source projection(s) and projected imagery of at least one reference 3D object.

The invention may further comprise the step of creating a 3D representation of the source 2D projection(s) by deforming the located (i.e., best-fitting) reference 3D representation so as to resemble the source multi featured object. In one embodiment, the deformation is a large deformation diffeomorphism, which serves to preserve the geometry and topology of the reference 3D representation. The deformation step may deform the located 3D representation so that feature items in the source 2D projection(s) align with corresponding features in the located reference 3D representation. The deformation step may occur with or without rigid motions and may include affine motions. Further, the deformation step may be constrained by at least one of known 3D position information associated with the source 2D projection(s), and 3D data of the source object. The deformation may be performed using a closed form expression.

In a second aspect, the invention comprises a system for comparing at least one source 2D projection of a source multifeatured object to a reference library of 3D reference objects. The system comprises a database comprising a plurality of reference 3D representations of generically similar multifeatured objects and an analyzer for performing a viewpoint-invariant search of the reference 3D representations to locate the reference 3 D representation having a 2D projection most resembling the source projection(s). In some embodiments, the analyzer determines resemblance by a degree of alignment between feature items in the 3D representation and corresponding feature items in the source 2D projection(s). The analyzer may search each reference 3D representation over a range of possible 2D projections of the 3D representation without actually generating any projections. In some embodiments, the analyzer searches over a range of possible 2D projections by computing a rigid motion of the reference 3D representation optimally consistent with a viewpoint of the source multi featured object in at least one of the 2D projections. The rigid motions may comprise pitch, roll, yaw, and translation in three dimensions. The analyzer may be configured to perform automatic camera calibration by estimating camera parameters, such as aspect ratio and field of view, from image landmarks.

In some embodiments, the analyzer is configured to determine the optimum rigid motion by estimating a conditional mean of feature items comprising points, curves, surfaces, and subvolumes in a 3D coordinate space associated with the reference 3D representation such that the feature items are projectionally consistent with feature items in the source 2D projection(s). The analyzer is further configured to generate MMSE estimates between the conditional mean estimate of the projected feature items and corresponding feature items of the reference 3D representation. The rigid motion may be constrained by known 3D position information associated with the source 2D projection(s).

In some embodiments, the analyzer is configured to extract feature items from the source projection using dynamic programming. In further embodiments, the analyzer may be configured to find feature items used in the matching process automatically by using correspondences between source imagery and projected imagery of at least one reference 3D object.
The invention may further comprise a deformation module for creating a 3D representation of the at least one source 2D projection by deforming the located (i.e., best-fitting) reference 3D representation so as to resemble the source multifeatured object. In one embodiment, the deformation module deforms the located reference 3D representation using large deformation diffeomorphism, which serves to preserve the geometry and topology of the reference 3D representation. The deformation module may deform the located 3D representation so that feature items in the source 2D projection(s) align with corresponding features in the located reference 3D representation. The deformation module mayor may not use rigid motions and may use affine motions. Further, the deformation module may be constrained by at least one of known 3D position information associated with the source 2D projection(s), and 3D data of the source object. The deformation module may operate in accordance with a closed form expression.

In a third aspect, the invention comprises a method of comparing a source 3D object to at least one reference 3D object. The method involves creating 2D representations of the source object and the reference object(s) and using projective geometry to characterize a correspondence between the source 3D object and a reference 3D object. For example, the correspondence may be characterized by a particular viewpoint for the 2D representation of the 3D source object.

In a fourth aspect, the invention comprises a system for comparing a source 3D object to at least one reference 3D object. The system comprises a projection module for creating 2D representations of the source object and the reference object(s) and an analyzer which uses projective geometry to characterize a correspondence between the source 3D object and a reference 3D object.

In a fifth aspect, the above described methods and systems are used for the case when the 3D object is a face and the reference 3D representations are avatars.

In a sixth aspect, the invention comprises a method for creating a 3D representation from at least one source 2D projection of a source multi featured object. In accordance with the method, at least one reference 3D representation of a generically similar object is provided, one of the provided representation(s) is located, and a 3D representation of the source 2D projection(s) is created by deforming the located reference representation in accordance with the source 2D projection(s) so as to resemble the source multi featured object. In some embodiments, the source 2D projection(s) is used to locate the reference representation. In further embodiments' the set of reference representations includes more than one member, and the reference most resembling the source 2D projection(s) is located by performing a viewpoint -invariant search of the set of reference representations, without necessarily actually generating any projections. The search may include computing a rigid motion of the reference representation optimally consistent with a viewpoint of the source multifeatured object in at least one of the source projections.

In a preferred embodiment, a 3D representation of the source projection(s) is created by deforming the located reference representation so as to resemble the source multifeatured object. The deformation may be a large deformation diffeomorphism. In some embodiments, the deformation deforms the located reference so that feature items in the source projection(s) align with corresponding feature items in the located 3D reference representation. In some embodiments, the deformation is performed in real time.

In a seventh aspect, the invention comprises a system for creating a 3D representation from at least one source 2D projection of a source multi featured object. The system includes a database of at least one reference 3D representation of a generically similar object, and an analyzer for locating one of the provided representation(s). The system further includes a deformation module for creating a 3D representation of the source 2D projection(s) by deforming the located reference representation in accordance with the source 2D projection(s) so as to resemble the source multi featured object. In some embodiments, the analyzer uses the source 2D projection(s) to locate the reference representation. In further embodiments, the set of reference representations includes more than one member, and the analyzer locates the reference most resembling the source 2D projection(s) by performing a viewpoint-invariant search of the set of reference representations, without necessarily actually generating any projections. The search may include computing a rigid motion of the reference representation optimally consistent with a viewpoint of the source multifeatured object in at least one of the source projections.

In a preferred embodiment, the deformation module creates a 3D representation of the source projection(s) by deforming the located reference representation so as to resemble the source multifeatured object. The deformation 45 may be a large deformation diffeomorphism. In some embodiments, the deformation module deforms the located reference so that feature items in the source projection(s) align with corresponding feature items in the located 3D reference representation. In some embodiments, the deformation module operates in real time.

VIEWPOINT-INVARIANT DETECTION AND IDENTIFICATION OF A THREE-DIMENSIONAL OBJECT FROM TWO-DIMENSIONAL IMAGERY –

US 7,853,085 B2

The present invention provides an automated method and system for identifying a 3D multifeatured object when only partial source information representing the object is available. Typically, the source information takes the form of one or more 2D projections of the 3D object, but may also include 3D data, such as from a 3D camera or scanner. The invention uses a set of candidate 3D representations of multifeatured objects, at least one of which is to be identified with the source object should a successful identification be made. In order to detect and locate the position of the source object in the source 2D image, the invention searches for feature points, curves, surfaces, or subvolumes which are characteristic of the 3D object and are substantially invariant under varying viewpoint and lighting. Next, all possible positions (i.e., orientations and translations) for each 3D candidate representation are searched to identify the candidate representation for which the optimal rigid motion (rotation and translation) has a projection which most closely matches the source feature items. The closeness of the match is determined by a measure such as the minimum mean-squared error (MMSE) between the feature items in the projection of the 3D representation and the corresponding feature items in the 2D source image. The comparison is performed in 3D between the estimated deprojected positions of the feature items from the 2D source image and the corresponding feature items of the candidate 3D representation. The rigid motion of the closest-fitting 3D candidate representation is tuned further by comparing portions of the source 2D imagery with corresponding portions of the projected best-fitting 3D representation. The quality of fit between the source imagery and the selected 3D representation in the tuned position is then determined. If the fit rises above a predetermined threshold, the identification is successful.

Thus the identification preferably proceeds in a hierarchical fashion. First, the presence of the object to be identified is detected in the source imagery using a coarse detection technique that rapidly locates a small number of feature items. Second, the candidate representations are searched for correspondence to these feature items across a continuum of possible viewpoints. Third, the optimal position of the best fitting candidate representation is refined by determining the best match between source imagery and projections of the 3D representation.
In a first aspect, therefore, the invention comprises a method of identifying a multifeatured object corresponding to at least one source 2D projection of a source

multi featured object. In accordance with the method, a set of candidate 3D representations is provided. The source 2D projection is detected within the source imagery and a viewpoint-invariant search of the candidate 3D representations is performed to locate the candidate 3 D representation having a 2D projection most resembling the source 2D projection(s). This candidate 3D representation is then compared with the source 2D projection(s) to determine whether the candidate corresponds to the source. In some embodiments, the detection of the source 2D projection is performed by locating viewpoint-invariant and lighting-invariant feature items in the source 2D image. This detection may be performed in real time. Each candidate 3D representation may be searched over a range of possible 2D projections of the 3D representation without actually generating any projections. The search over a range of possible 2D projections may comprise computing a rigid motion of the candidate 3D representation optimally consistent with a viewpoint of the source multifeatured object in at least one of the 2D projections.
In some embodiments, the optimum rigid motion is determined by estimating a conditional mean pose or geometric registration as it relates to feature items comprising points, curves, surfaces, and subvolumes in a 3D coordinate space associated with the candidate 3D representation such that the feature items are projectionally consistent with feature items in source 2D projection(s). MMSE estimates between the conditional mean estimate of the projected feature items and corresponding feature items of the candidate 3D representation are generated. A quality of fit is determined by, for example, comparing portions of the projection of the candidate 3D representation yielding the lowest MMSE estimate and corresponding portions of the source 2D representation, and determining the optimum rigid motion as corresponding to the best quality of fit. If the quality of fit corresponding to the optimum rigid motion exceeds a predetermined threshold, the candidate 3D representation is positively identified with the source object.

In a second aspect, the invention comprises a system for identifying a multifeatured object corresponding to at least one source 2D projection of a source multifeatured object. The system comprises a database comprising a plurality of candidate 3D representations of multifeatured objects, a detection module for detecting the source 2D projection in an image, and an analyzer. The analyzer performs a viewpoint invariant search of the candidate 3D representations to locate the one having a 2D projection most resembling the source 2D projection(s), and compares the source 2D projection(s) with the located candidate 3D representation to determine whether the candidate corresponds to the source. In some embodiments, the detection module detects the source 2D projection by locating viewpoint-invariant and lighting-invariant feature items in the source 2D image. This detection may be performed in real time. The analyzer may search each candidate 3D representation over a range of possible 2D projections of the 3D representation without actually generating any projections. The search over a range of possible 2D projections may comprise computing a rigid motion of the candidate 3D representation optimally consistent with a viewpoint of the source multifeatured object in at least one of the 2D projections.

In some embodiments, the analyzer determines the optimum rigid motion by estimating a conditional mean pose or geometric registration as it relates to feature items comprising points, curves, surfaces, and subvolumes in a 3D coordinate space associated with the candidate 3D representation such that the feature items are projectionally consistent with feature items in source 2D projection(s). The analyzer generates MMSE estimates between the conditional mean estimate of the projected feature items and corresponding feature items of the candidate 3D representation. The analyzer further generates a quality of fit by, for example, comparing portions of the projection of the candidate 3D representation yielding the lowest MMSE estimate and corresponding portions of the source 2D representation, and determines the optimum rigid motion as corresponding to the best quality of fit. If the quality of fit corresponding to the optimum rigid motion exceeds a predetermined threshold, the analyzer positively identifies the candidate 3D representation with the source object.

In a third aspect, the above described methods and systems are used for the case when the 3D multifeatured object is a face, and the candidate 3D representations are avatars.