NUI Galway
OÉ Gaillimh

ARAN - Access to Research at NUI Galway

Provided by the author(s) and NUI Galway in accordance with publisher policies. Please cite the published version when available.

Title	In-camera person-indexing of digital images
Author(s)	Costache, G.; Mulryan, R.; Steinberg, E.; Corcoran, Peter
Publication Date	2006-01-07
Publication Information	Costache, G., Mulryan, R., Steinberg, E., & Corcoran, P. (2006, 7-11 Jan. 2006). In-camera person-indexing of digital images. Paper presented at the 2006 Digest of Technical Papers International Conference on Consumer Electronics, Las Vegas, NV, USA, 07-11 January.
Publisher	Institute of Electrical and Electronics Engineers
Link to publisher's version	https://dx.doi.org/10.1109/ICCE.2006.1598449
Item record	http://hdl.handle.net/10379/16690
DOI	http://dx.doi.org/10.1109/ICCE.2006.1598449

Downloaded 2022-08-10T05:17:46Z

Some rights reserved. For more information, please see the item record link above.

BY NC ND

In-Camera Person-Indexing of Digital Images

Costache, G, Mulryan, R, Steinberg, E, and Corcoran, P, *Member, IEEE*

***Abstract*--A prototype implementation of an automatic in-camera cataloging tool is presented. An external infrastructure to store and analyze images and support the in-camera cataloging tool is also described..**

I. Introduction

Digital cameras continue to be one of the CE industry's recent success stories. However as many users switch from conventional to digital photography they find themselves with rapidly growing collections of digital images. Few consumers have the time and personal discipline to manually catalog and organize their personal image collections.

In earlier publications [1, 2] we presented an automatic cataloging tool for consumer image collections. This tool employed a combination of face detection, face recognition, face region normalization and color and texture-based analysis tools to allow automated annotation of an image collection. One of the principle challenges in implementing a workable cataloging system is the ad-hoc nature of consumer images. For example the quality of identified face regions is often poor and unsuitable for conventional face recognition techniques. Furthermore, our system allows a persons clothing, hair and other distinctive features to be incorporated as distinguishing aspects, in addition to more traditional techniques such as face recognition.

In this paper we present a refinement of the original prototype system which is intended to support the cataloging of images within a digital camera as they are actually captured. This requires modifications to the main system architecture and a separation of the application into an embedded systems component and a desktop/server based component. The principle goal of this design is to lowers the computational and storage requirements of the in-camera application components by off-loading as many computational and storage tasks to the desktop/server infrastructure. In addition our system incorporate specialized techniques to minimize and optimize initial training and the required re-training necessary to incorporate data from new image sets..

II. System Architecture

The main system architecture is illustrated in ***Figs 1(a)*** and ***1(b)*** below.

Fig 1(a) shows the overall architecture of the image classification system. An underlying classification engine performs an analysis of new images and stores the images and associated image data which can be determined automatically. A main workflow module facilitates user interaction with the image collection. This enables more advanced image analysis and annotation functionality which requires user interaction.

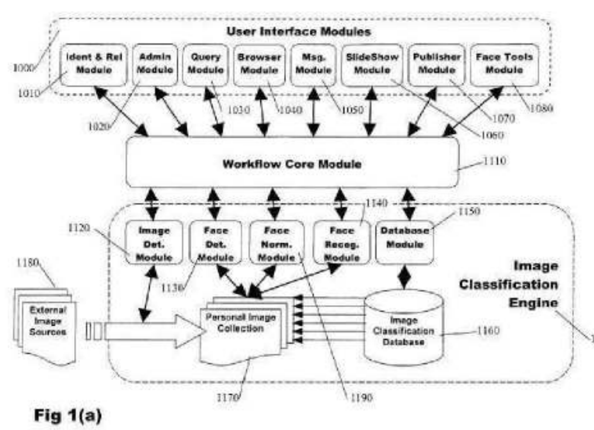

Fig 1(a): Image Classification Engine and UI Components.

Fig 1(b) illustrates how this is modified to separate the system into an embedded component for incorporation within the firmware of a digital camera, and a desktop/server component which maintains the main database for the image collection.

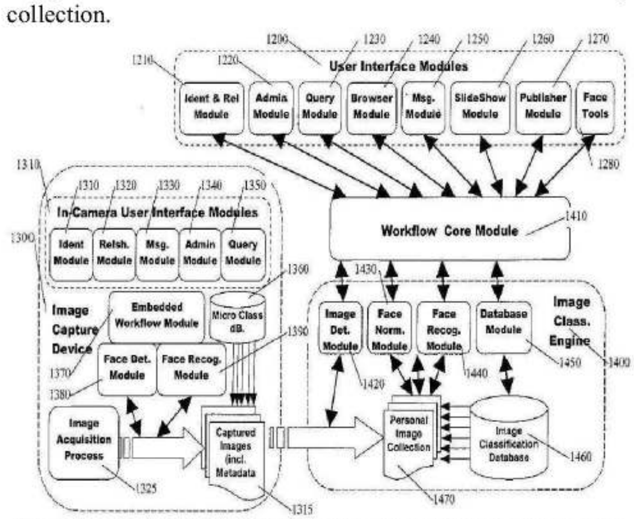

Fig 1(b): Separation of System Functionality into Embedded System and Desktop/Server Components.

Given that many digital cameras now come equipped with "docking stations" this form of distributed architecture is very much in keeping with current trends in digital camera design. We also remark that some of the more recent models of digital camera have begun to incorporate face detection capabilities and the current state-of-art of image processing capabilities for high-end digital cameras can readily support more sophisticated algorithms, particularly when they can be implemented as background image post-processing tasks outside of the main image acquisition chain.

By enabling in-camera classification and annotation of images a whole new set of practical user tools for image management, based on the people in a localized (in-camera) image set, becomes practical. As examples: images can be pre-sorted on the camera to allow sharing a mixed set of images with different groupings of family and friends; events or user notifications can be triggered in response to certain persons, or combinations of people; real-time statistical analysis can indicate that you are shooting too many/not enough images of certain individuals.

III. Image Classification Database

Fig 2 illustrates the internal structure of the main image classification database.

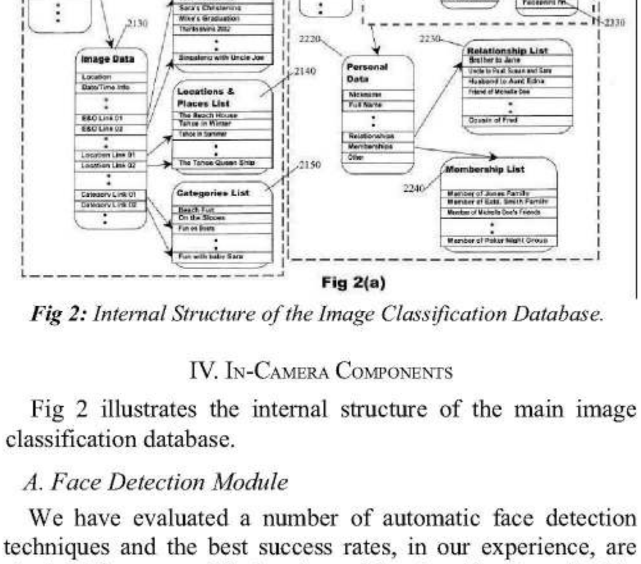

Fig 2(a)

Fig 2: *Internal Structure of the Image Classification Database.*

IV. In-Camera Components

Fig 2 illustrates the internal structure of the main image classification database.

A. Face Detection Module

We have evaluated a number of automatic face detection techniques and the best success rates, in our experience, are obtained from a modified variant of the algorithm described in [3]. This is also one of the fastest algorithms described in the literature. However it is worthwhile remarking that 90%+ of the time required by our original person-detection algorithm, described in [1, 2] was spent in the face detection phase.

Thus the algorithm described in [3] is not suitable for in-camera implementation in a conventional camera. Fortunately it is possible to speed up this algorithm and, by employing a combination of other image segmentation techniques and by confining our analysis to the central region of an image it is possible to achieve real-time face detection for in-camera applications.

B.Face Recognition Module

As we are generally restricted to relatively small subsets of people in a consumer image collection standard recognition algorithms can be employed with a high degree of success. Evaluations are presented of applying two methods for feature extraction: DCT and PCA *(principle component analysis)* on wavelet decomposition [4, 5, 6] and also a combination of the two [2]. Once training data are known for these techniques the actual thresholding and comparison of face regions can be implemented quite efficiently and are very suitable for in-camera applications. Image sorting can also be improved by identifying particular items of clothing & jewelery and by analyzing the color and texture of a person's hairstyle [2]. These supplemental analysis techniques provide an additional method to sort and categorize images and, as the location of these regions is well-defined in relation to detected face regions it is straightforward to add this functionality to the main face recognition module.

V.System Workflow

Training is performed ex-camera using the main image collection. In our original algorithm [2] many of the classifiers are absolute and thus do not require retraining. However the PCA components used for face recognition are dependent on the main image collection. Thus occasional retraining is required. Fortunately we have developed techniques which do not require that the PCA classifiers are retrained across the entire image collection. Thus it is only necessary to retrain using newly acquired images and the training time required does not increase as the size of a user's collection grows..

Thus, when the camera is docked the latest image set is off-loaded to the main classification system. Most images will already have marked face regions and an automatic retraining process can be started. Some images may require additional user input which can be flagged to the user after the image set is off-loaded from the camera.

After training the collection the modified recognition classifiers are loaded back onto the camera and will be applied when the next set of images is shot by the user.

Acknowledgment

.This research was funded under the *Technology Development Phase* of the Enterprise-Ireland Commercialization Fund.

Examples of Reference Styles

[1] P. Corcoran and G. Costache, "Automated Sorting of Consumer Image Collections using Person Recognition Techniques", *IEEE International Conference on Consumer Electronics*, Las Vegas, *5,* Jan 2005; pp127-128.

[2] P. Corcoran and G. Costache, "Automated Sorting of Consumer Image Collections using Face and Peripheral Region Image Classifiers", *IEEE Trans. Consumer Electronics,* Volume 51, Issue 3, pp 747 – 754, August 2005 .

[3] P. Viola, M. J. Jones "Rapid Object Detection using a Boosted Cascade of Simple Features". IEEE CVPR, pp. 511-518, 2001.

[4] M.Turk, A.Pentland, "Eigenfaces for recognition" Journal of Cognitive Neuroscience, 3 (1), 1991

[5] D. Beymer and T. Poggio "Face recognition from one example view" Proceedings of the IEEE 5th International Conference on Computer Vision, IEEE Computer Society Press, Cambridge, MA, 500-507, June 1995.

[6] R. Tjahyadi, W. Liu, S. Venkatesh "Application of the DCT energy histogram for face recognition" Proceedings of the 2nd International Conference on Information Technology for Application (ICITA 2004)