



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

Title	FPGA based stereo imaging system with applications in computer gaming
Author(s)	Andorko, Istvan; Corcoran, Peter
Publication Date	2009-10-23
Publication Information	Andorko, I., Corcoran, P. M., & Bigioi, P. (2009). FPGA based stereo imaging system with applications in computer gaming. Paper presented at the Games Innovations Conference, 2009. ICE-GIC 2009.
Publisher	IEEE
Link to publisher's version	http://dx.doi.org/10.1109/ICEGIC.2009.5293586
Item record	http://hdl.handle.net/10379/1355

Downloaded 2024-03-20T12:11:09Z

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



FPGA Based Stereo Imaging System with Applications in Computer Gaming

Istvan Andorko and
Peter M. Corcoran, *Senior Member IEEE*
College of Engineering and Informatics,
National University of Ireland Galway, Ireland
i.andorko1@nuigalway.ie, peter.corcoran@nuigalway.ie

Petronel Bigioi, *Member IEEE*
Tessera (Ireland) Ltd,
Parkmore Industrial Estate,
Galway, Ireland
pbigioi@tessera.com

Abstract—A real-time stereo imaging system is described. It incorporates dual image acquisition chains on a single FPGA device and is able to provide real-time synchronized video output from twin CMOS imaging sensors. The platform is designed to provide hardware support for future implementations of advanced face analysis algorithms. In the context of computer gaming applications this system can provide real-time capture of a game players facial actions and expressions serving as an enabling technology for next generation gaming. Potential applications include (i) generating 3D real-time responsive game avatars and (ii) employing real-time face data for next-generation game UI.

Keywords- stereo-imaging, imaging pipeline, face detection & analysis, game peripheral

I. INTRODUCTION

Face detection and tracking technology has become commonplace in digital cameras in the last year or so. Most practical embodiments of this technology are based on haar classifiers and follow some variant of the classifier cascade originally proposed by Viola and Jones [1]. These haar classifiers are rectangular and by computing a grayscale integral image mapping of the original image it is possible to implement a highly efficient multi-classifier cascade. These techniques are well suited for hardware implementations [2].

However current techniques can only determine the approximate face region and do not permit any detailed matching to facial orientation or pose. A more accurate matching of such characteristics of the face region enables more sophisticated use of face modeling techniques. Applications which are relevant to computer gaming include (i) the detection and recognition of game-players both for authentication and to enable personalization of the gaming environment; (ii) the detection and modeling of user facial expressions which can be employed to provide feedback to the gaming environment on the mood and frame of mind of players, (iii) the detection and tracking of facial expressions to enable real-time animation of gaming avatars. While we will not deal explicitly with any of these topics in this paper they form the basis for our motivation to develop our FPGA stereo imaging peripheral. In companion papers we will explore aspects of facial expression modeling [3] and stereoscopic AAM face models [4]. The imaging peripheral described in this paper serves as an enabling technology for these face modeling and analysis techniques.

One of the biggest challenges in real-time face modeling is the determination of a 3D representation of the face region. Conventional capture of the facial image by a single imaging device only permits matching of the face to 2D templates or models. To obtain a fully accurate representation would ideally require at least a stereo image of the face using two conventional cameras with different physical perspectives on the face region. However a low-cost consumer electronics solution to this problem has not been available restricting the practical development of gaming techniques employing user face-feedback.

In this paper we present our work in developing such a low-cost stereo imaging solution based on a pair of conventional CMOS imaging sensors and a single FPGA device incorporating a dual synchronized image pipeline. Additional post-processing of the dual image streams is described in two companion papers..

II. BACKGROUND AND RATIONALE

In this section we provide additional background information so the reader can better understand the context for our stereoscopic imaging platform. In this regard we begin by reviewing the existing DSP and FPGA based solutions which provide hardware platforms for advanced face detection and analysis in real time. A discussion of some advanced face detection and analysis methods is then provided. Finally we summarize some of the existing work on active appearance face models (AAM). We provide initial arguments why this family of face models is the most effective for our imaging platform.

A. DSP Implementations of Facial Analysis Algorithms

There are two DSP-based *automated facial recognition* (AFR) system implementations by Batur, Flinchbaugh, and Hayes [6] and Wei and Bigdeli [7] that are reported in the literature.

The implementation of [6].consists of four stages: (i) face detection, (ii) face feature localization, (iii) face normalization, and (iv) face recognition. The probabilistic visual learning algorithm proposed by Moghaddam and Pentland [8] is used for face detection. The system was implemented on the Texas Instruments TMS320C6416 fixed point DSP operating at 500 MHz. On the other hand, the system implemented by Wei and Bigdeli [7] is made up of three stages: (i) image normalization, (ii) face detection, and (iii) face recognition. Where face

detection is performed using the algorithm proposed by Rowley, et al. [9]. The system was implemented on the Analog Devices ADSP-BF535 EZ-KIT Lite development board containing a 16-bit fixed point DSP operating at 300MHz.

Similar optimization techniques were employed by both systems, including converting floating point operations to fixed point, write time-consuming functions in assembly, use of available parallelism in the DSPs, and use of look-up tables in place of complex arithmetic operations. Unfortunately, both implementations still require significantly more than one second to process an image with face detection consuming the majority of the processing time. Such DSP based approaches may be improved with more modern devices but a speed up of two orders of magnitude is needed for practical real-time face analysis to be implemented using DSP based systems at typical video frame rates of 30 to 60 frames per second.

B. FPGA Implementations of Facial Analysis Algorithms

In applying any of these modeling techniques it is important to achieve as accurate an initial alignment of the model with the detected face region. For embedded systems it is important to be able to apply and iterate models at high speeds which is often a challenge for the available processing power. While software implementations can run at adequate frame rate on a desktop computer it becomes challenging to implement complex image processing algorithms on an embedded system.

A number of other authors have looked to FPGA devices to provide a pragmatic means to implement sophisticated face analysis and detection techniques. C Gao and S. L. Lu present an approach to use FPGA to accelerate the Haar-classifier based face detection algorithm [10]. They have shown that for the software version, the Haar classifier face detection application was only able to achieve a performance of 5 frames/sec, whereas for 1-classifier FPGA implementation 37 frames/sec and for 16-classifier FPGA implementation 98 frames/sec.

In their paper, I. Sajid et al. present an FPGA based design developed for efficient face recognition system which provides software hardware co-design, customization of algorithm and adaptability in the system [11]. They have shown that their proposed system is reasonably power efficient than floating point architecture and can be employed for portable applications.

Y. Wei et al. present a FPGA design of hardware for a real-time face detection based on AdaBoost algorithm [12]. They state that being fully programmable, hardware design using FPGA offers a much shorter development time and enables a quick verification of DSP algorithms. Their results have shown that based on the synthesis result, the FPGA can operate at 15 frames/sec which offers the computation performance of a high-end PC at very low cost.

C. Automated Face Analysis Systems

As many readers may not be familiar with this field it seems appropriate to provide a succinct introduction to the concepts of automated facial analysis systems. In Figure 1 we illustrate the key elements of such a systems. After image or video acquisition the first step is that of *face detection*. This non-trivial task is often the most expensive both in terms of

time required and of processing power. For video we can reduce the burden somewhat by re-use of information from frame-to-frame and this represents a form of refinement known as *face tracking*. Most modern camera algorithms use tracking enhancements applied to the *preview video stream* – the real-time stream of images that is provided on the display of a digital camera in lieu of the traditional eyepiece.

After one or more confirmed face regions have been detected these are subject to further *intermediate* analysis in order to provide more detailed information about the face region. Among the more commonly used types of analysis we find determinations of facial *sharpness*, *pose*, *in-plane rotation*, *eye-state*, *mouth-state*, *skin tone*, *color* and *illumination histograms*, *occlusions* (e.g. *glasses or hair*) and *facial boundary*. More sophisticated analysis can provide additional details about *eye-gaze* and *facial expressions*. Ultimately all of these methods rely on some underlying determination of *facial features* extracted from the detected face region(s). There are many different approaches to facial feature analysis in the literature, many being influenced by the ultimate goals of a particular facial analysis application. In a video sequence a statistical determination of these features is possible.

In many cases an explicit *face normalization* step may be used prior to feature extraction. As the most frequent sources of error are due to facial *pose* and *illumination* this normalization step is frequently directed to a correction of one or both of these aspects of the detected face region(s). Note that often this step may form a part of the actual feature extraction process. For example, in many facial models the largest variations are due to *pose* and *illumination* thus the lowest order components of the face model will often incorporate these aspects of the face. Thus it may be sufficient simply to disregard these low-order model components and no explicit *normalization* step is necessary. .

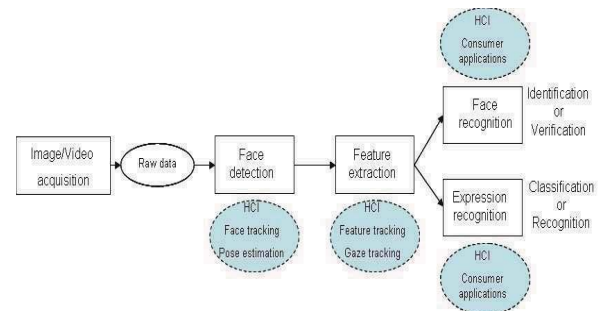


Figure 1: Generic Facial Analysis/Classification System

Following this *feature detection* step we have the main application. This can range from *facial autofocus* in a digital camera to a determination of *facial expression* for medical or security applications. Probably the best known end-application for face analysis is that of *automated facial recognition* which has received much attention in recent years as a means of improving security at airports. Again there are many differing approaches to *facial recognition* and each technique tends to have a preferred method of feature extraction to meet its needs.

Perhaps the best know technique for face recognition is that of *Eigenfaces*, originally proposed by Turk and Pentland [5]. Their approach leads to a sequence of 2D eigenvectors where

the lower order components have the global shape of the face and higher order components capture more finely textured aspects of the face region. As the first few components capture the global shape of a face regions these eigenvectors are known colloquially as *eigenfaces*.



Figure 2: Examples of lower order eigenface Vectors

D. Active Appearance Face Modeling Techniques

One particular class of face models we are interested in are known as Active Appearance Models [13], [14]. These are affine 2D models of a face region which use separate shape and texture subspaces to model the face region. There are a multitude of alternative applications for such models. These models have been widely used for face tracking [10], and measuring facial pose and orientation [15], [16]. They can also be used to “repair” facial regions. Consider, for example, a face region that has an eye defect, or perhaps a blemish or food stain on the face. By applying an AAM we can obtain a best fit to the face region based on the texture model employed in the training set. Now if the face has been disfigured by a defect the modeled face will not be able to reconstruct the defect area if it was not provided in the training set, but it will be able to accurately reconstruct the global face region. Portions of this reconstructed face region can then be substituted for defects in the original image effectively repairing it. An extension of this technique has been proposed as a means to remove global defects such as acne [17]. Another variation has been proposed to filter periodic noise components from the facial regions of an image [18].

In other research we have demonstrated the use of AAMs for detecting phenomena such as eye-blink [19], analysis and characterization of mouth regions [20] and facial expressions [21]. In such context these models are more sophisticated than other pattern recognition methods which can only take a binary decision that an eye is in either an “open” or a “closed” state. Our models can determine other metrics such as the degree to which an eye region is open or closed or the gaze direction of the eye [22].

E. End Goals of our Research

Finally we should explain our broader motivations. In implementing a stereo imaging platform within an FPGA system we have a number of end goals which extend beyond the results presented in this paper. Our principle motivation is actually to construct a hardware platform which enables multiple parallel image acquisitions to be effected at the same time. The broader vision behind this effort is the availability of low cost high-quality wafer level cameras [39] will drive a range of new imaging applications where multiple images can be acquired of a single scene under different exposure, focus and timing criteria. Challenges exist in the synchronization and timing of such multiple acquisitions but by overcoming these we hope to present a new hardware platform which can act as an enabling technology for the development of applications based on synchronized multi-image acquisition. Thus our work

in stereoscopic imaging represents a first step towards this end goal.

III. STEREOSCOPIC IMAGING SYSTEM OVERVIEW

This section explains the fundamentals of the stereo imaging system. We begin with a review of the literature on stereo imaging techniques and applications. This is followed by an overview of the hardware architecture of our system, a discussion of the hardware/software integration and an overview of the system operation.

Finally we discuss testing and evaluation of the system and give details of some of the practical problems that were encountered during our initial development, including some EMI effects which led to a noticeable physical offset of one of the stereo image pairs. In the end the system functioned well and we have used output images from this system to provide stereo image pairs for developing related methods to create 3D face depth maps and even rudimentary 3D facials models [4].

A. Stereo Imaging Techniques and Applications

Most of the relevant literature on stereoscopic imaging describes applications which use a stereo image processing pipeline (IPP). We did not find many papers in the literature devoted to practical implementation of such a pipeline. We first consider research papers which present filtering and image compression methods for stereo image streams.

S-H Seo et al. [23] presents a two-dimensional least squares based filtering scheme for high fidelity stereo image compression applications. This method removes the effect of mismatching in a stereo image pair by applying the left image as the reference input to a 2D transversal filter while the right image is used as the desired output. M. Moellenhoff and M. W. Maier [24] present image compression techniques specific to stereo imaging and compare performance with non-stereo methods. Image compression techniques can be utilized to reduce the transmission bandwidth and/or storage space requirements of the stereo pairs. F. Davoine et al. [25] present a scheme for fractal image compression based on adaptive Delaunay triangulation. We remark that while stereo image compression would be important for a commercial system we did not concern ourselves with non-standard compression and much of our initial work has been directed to working with uncompressed images.

Another important aspect of stereo images is how information in the image pairs can be used to interpolate or extrapolate data. In this regard J. D. Boissonnat et al. [26] propose a coherent way of interpolating 3D data from 2D stereo image pairs. Their proposed method is based on the use of constrained Delaunay triangulation: a polyhedral surface is obtained by using a simple visibility property to mark tetrahedra likely to be empty. In [27] W-J Kim and J. Kim present a frequency analysis of stereoscopic 3D image and the corresponding development of an anti-aliasing filter. The frequency characteristics of 3D images are analyzed using a geometry model, and it is confirmed that there is an inherent aliasing of data due to negative disparity which is eliminated through use of the anti-aliasing filter.

The next papers describe some applications which take advantage of stereoscopic imaging and give some background

as to the potential of stereoscopic imaging systems. In [28] S. Takezawa and G. Dissanayake propose a method for simultaneous localization and mapping (SLAM) in an indoor environment using stereo vision. In this method specially designed artificial landmarks distributed in the environment are observed and extracted from a camera image. The disparity map obtained from the stereo vision system is used to obtain the ranges to these landmarks. In [29] G. Roth presents a method of computing camera positions from a sequence of overlapping images obtained from a binocular/trinocular camera head. This method relies on finding matching features among the images at each camera head position and after this, directly computing the 3D coordinates of these features using triangulation. In [30] by A. Baumberg a robust method for automatically matching features in images corresponding to the same physical point on an object seen from two arbitrary viewpoints is presented. In this method, features are detected in two or more images and characterized using affine texture invariants. The feature matching process is optimized for a structure-from-motion application where unreliable matches are ignored at the expense of reducing the number of feature matches.

Based on our review of the literature we did not find any practical hardware implementation of an integrated stereo imaging device. Most of the stereo image acquisition pipelines are implemented in software on COTS computer systems, or as two discrete hardware pipelines without explicit synchronization mechanisms between them. An exception is the work of Ben-Ezra et al. [40], [41] and the related work of Yu-Wing Tai et. al. [42] where multiple cameras are used to capture fast low-resolution images and corresponding high resolution images of a scene. The authors of [42] use a pair of stereo high-speed, low-resolution cameras to compliment the single high resolution camera. The work in these papers will provide the interested reader with some ideas of what is possible with synchronized multi-image acquisition systems and further explain our interest in creating an FPGA-based enabling platform for further experimentation with low-cost consumer imaging sensors. In the context of this paper, not only does an FPGA based system offer much greater flexibility and cost savings, but also enables a practical, low-cost computer peripheral to be realized. If we can incorporate additional hardware based image processing such as face detection and modeling we can provide a useful peripheral to enable advanced functionality for modern game consoles, drawing the user more intimately into the gaming experience. We will explore these ideas in more detail later in this paper.

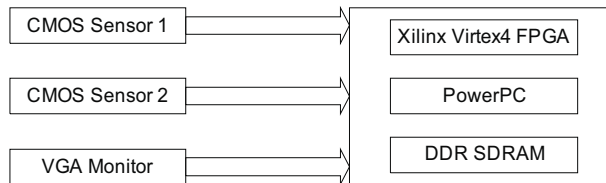


Figure 3: General Architecture

B. Hardware Architecture of Stereo Imaging System

The general architecture of the system is illustrated in Figure 3 above. Note that the use of independent external imaging sensors enables greater flexibility in deciding their relative positioning and also enables the use of twin sensors

with differing capabilities, e.g. one wide angle sensor, combined with a normal field of view sensor. While we will not explore such heterogeneous imaging in this paper it is worthwhile noting that it is possible with our system.

The internal architecture of the design is detailed in the following. Figure 4. The main image data is carried over the processor local bus (PLB) while control signals are passed between the subsystems over the DCR bus.

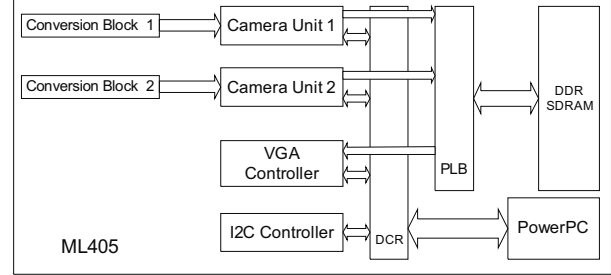


Figure 4: Internal Architecture

The development board is a Xilinx ML405 development board, with a Virtex 4 FPGA, a 64 MB DDR SDRAM memory, and a PowerPC RISC processor. The clock frequency of the design is 100 MHz.

The sensors used are 1/3 inch SXGA CMOS sensors made by Micron. They have an active zone of 1280x1024 pixels. They are programmable through an I2C interface. They typically run at 13.9 fps and a clock frequency of 25 MHz [31]. This sensor was selected because of its small size, low cost and the imaging capabilities are matched to the memory size and processing bandwidth of the Virtex4 FPGA. This system enables real-time stereo video capture at a standard VGA resolution of 640x480 with a fixed distance between the two imaging sensors.

C. Hardware-Software Interface

The Hardware design is completely controllable from the RISC PowerPC microprocessor. Elements that can be controlled by the processor are the two sensors, through the I2C bus [32], the down sampling scale on the acquired image, the base addresses which are used to write and read to/from the memory. There are ways to communicate between the processor and the hardware design in both directions. The RISC processor addresses the registers and memory in Big Endian style whereas the hardware design addresses the registers in Little Endian style [33]. Care has to be taken when data is being sent from the software to the hardware or the other way around. This possibility of communication is essential, because, based on the processing made in hardware, certain flags can be set, which can be later used in the processor for certain settings of the system.

D. System Operation

The first operation that is done is to set the working parameters of the sensors through the I2C bus. The resolution of the sensors has to be set to 640x480 because this way there is less data to work with and the gain values of the sensors for the white balance are set to the appropriate values [31]. The data coming from the sensors is in Bayer format [34]. The first step is to convert the data from Bayer format to RGB format

[34]. This operation is performed in parallel by both acquisition chains. The importance of this operation is significant because the image processing operations are easier to implement for images available in RGB format.

After this step, data is sent to the camera unit, which bundles the data into 64 bit registers and requests access to the Processor Local Bus (PLB) which is the connection between the camera unit and the DDR SDRAM memory. Because of the two different acquisition chains, there has to be an arbitration module, which allows them to connect to the PLB bus one at a time, in a Round Robin manner. The same PLB bus is used for the connection between the VGA Controller and the DDR SDRAM as well. The VGA Controller reads the data from the DDR SDRAM, unbundles the 64 bit register into individual 8 bit registers for the R, G and B color components, generates the synchronization signals and sends the data to the monitor together with the synchronization signals [35].

E. Simulation and Testing the FPGA System

1) VeriLog Device Utilization and Timing reports

For the testing and simulation of the design, a Modelsim PE 6.3 hardware simulation software has been used. Instead of the CMOS sensor, a VerilogHDL model of the MT9M011 Sensor has been used, which had as an input an image converted from RAW format to RGB format. The VerilogHDL model of the sensor converted the RGB file into Bayer format, and this way we were able to test the conversion module of our design as well. To be sure that everything was working right, dump of data into some files had to be made, which later was checked with the IrfanView image processing software. Figure 5 represents the architecture of the testing system using the simulation software.

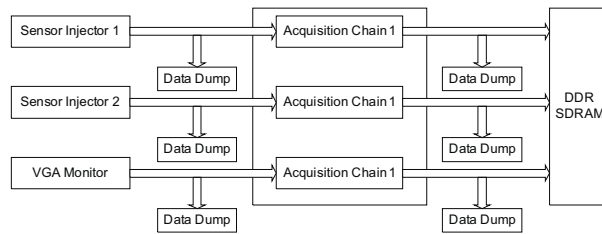


Figure 5: Architecture of Testing System for Dual Sensor Design

For the testing of the design in real-time, a Xilinx ML405 development board, two MT9M011 CMOS sensors and a CRT monitor have been used. This test proves the correct functionality of the system.

General			
IP Core	dual_pipeline		
Version	1.00.a		
Post Synthesis Device Utilization			
Resource Type	Used	Available	Percent
Slices	1268	8544	14
Slice Flip Flops	1134	17088	6
4 input LUTs	1982	17088	11
IOs	1303	NA	NA
bonded IOBs	0	320	0
TBUFs	2	0	4
DSP48s	10	32	31

Figure 6: Device Utilization Report

For the interconnection of the different modules in the system, Xilinx Platform Studio development software has been used. It has generated the following reports about our IP. The device utilization report, which indicates how much space of the FPGA our IP occupies, Figure 6.

Figure 7 shows an example timing report for our IP. This report shows the maximum frequency at which the IP can operate, based on the delays in the design. These test results were further confirmed by bench testing of the final design. The maximum frame rate for these types of sensors at full resolution is 13.9 fps. For the resolution of 640x480 that it is used in our design, the frame rate is 25 fps. This frame rate applies to the stereo-pair frame rate. If we would like to increase the resolution of the sensors and in the same time the frame rate, the bottleneck will eventually be our design, because as the clock frequency of the sensors increases, the design won't have enough resources to deal with both pipelines simultaneously. There is the possibility to save the stereo image pairs on a compact flash card that is supported by the ML405 development board. For the debugging of the design in real-time we have used Chipscope Pro Analyzer software which allowed us to see that values of the registers and signals of the design in real time. The frequency of the vertical and horizontal signals coming from the sensor was analyzed with the help of a digital oscilloscope.

Post Synthesis Clock Limits		
These are the post synthesis clock frequencies. The critical frequencies are marked with green. The values reported here are post synthesis estimates calculated for each individual module. These values will change after place and route is performed on the entire system.		
MODULE	CLK Port	MAX FREQ
dual_pipeline_0	clk_100	111.562MHz
dual_pipeline_0	cam_xclk	111.562MHz
dual_pipeline_0	vga_clk	111.562MHz

Figure 7: Timing Report

2) Electromagnetic Incompatibility Problem

For the testing of the synchronised operation of the sensors, a test design was created, where at one of the sensors instead of using its own incoming synchronization signals, the ones from the other sensor were used. This is where the Electromagnetic incompatibility problem was encountered and example images can be found in Figure 8, where you can clearly see in the right picture that the frame is moved to the right. The reason for the offset in the image was that the data lines were delayed because of the EMI problem. We have made the assumption that the horizontal synchronization signal was influenced by the power supply cable.

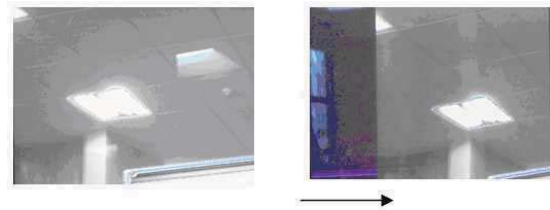


Figure 8: Electromagnetic Incompatibility Problem

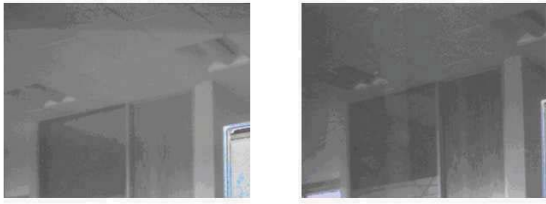


Figure 9: The EMI problem is solved

The EMI problem was solved by rebuilding the connections between the sensors and the board making the connections shorter, and connecting the analogue ground pin very close to the Vcc pin. Example images can be found in Figure 9, where you can clearly see that the frames are at the same position.

IV. IMAGING PLATFORM GAMING APPLICATIONS

Again we remind the reader that our principle motivation is to construct a hardware platform which enables multiple parallel image acquisitions to be effected at the same time. Challenges exist in the synchronization and timing of such multiple image acquisitions but by overcoming these we can present a new hardware platform which can act as an enabling technology for the development of applications based on synchronized multi-image acquisition. In this section we outline further details of some of the gaming related applications for our stereo imaging system.

A. Real-time avatars

Y. Fu et al have presented a novel framework of multimodal human-machine or human-human interaction via real-time humanoid avatar communication for real-world mobile applications [36]. Their application is based on a face detector and a face tracker. The face of the user is detected and the movement of the head is tracked detecting the different angles, sending these movements to the 3D avatar. This avatar is used for low-bit rate virtual communication. The drawback of this approach is that the shape of the avatar needs to be specified by the user and the forward-backward movement is not detected.

In a companion paper [4] we describe an enhanced face model derived from active appearance model (AAM) techniques which employs a differential spatial subspace to provide an enhanced real-time depth map. Employing techniques from advanced AAM face model generation [37] and the information available from an enhanced depth map we can generate a real-time 3D face model. The next step, based on the 3D face model is to generate a 3D avatar that can mimic the face of a user in real time. We are currently exploring various approaches to implement such a system using our real-time stereoscopic imaging system.

B. Advanced Gaming UIs

Our system can also be used to realize novel real-time user-interface methods. Again, by adding our system to a game console and observing a user we can make a real-time determination of their facial movements. Using depth map techniques we can also track forwards and backwards movements with a good precision which was not practical with simple 2D imaging.

One example of a simple UI mode based on this capability is to link player movements in a game environment to the relative position and movement of a player's head. Thus when the player leans forward their character or avatar will move forward within the game environment and when they react to a stimulus in the game, withdrawing backwards, or leaning left or right then the game character can mimic these movements. We are currently exploring some test environments to make further studies on the practical aspects of such a user-dynamic UI.

C. Full User Visualization (FUV)

The final step will be to combine real-time avatar modeling with user tracking to enable FUV. The fully virtualized user can then participate in multi-player role playing games as a real-time representation of themselves. This new *super-avatar* could be used for human-machine or human-human applications, where instead of sending real-time images of the user, the avatar will take the place of these images and can interact in a computer generated environment with other super-avatars.

ACKNOWLEDGMENT

This project has been part-funded by the Irish Research Council for Science, Engineering and Technology (IRCSET) and Tessera (Ireland) Ltd. .

REFERENCES

- [1] P. A. Viola, M. J. Jones, "Robust real-time face detection", *International Journal of Computer Vision*, vol.57, pp.137-154, 2004.
- [2] A. Bigdeli, C. Sim, M. Biglari-Abhari and B. C. Lovell, "Face Detection on Embedded Systems", *Proceedings of the 3rd international conference on Embedded Software and Systems*, Springer Lecture Notes in Computer Science, vol. 4523, pp. 295-308, May 2007.
- [3] I. Bacivarov and P. Corcoran, "Facial Expression Modeling using Component AAM Models – Gaming Applications", at *International IEEE Consumer Electronics Society's Games Innovations Conference 2009 (ICE-GIC 09)*, London, UK.
- [4] M.C. Ionita, I. Androkov and P. Corcoran, "Enhanced Real-Time Face Models from Stereo Imaging for Gaming Applications", at *International IEEE Consumer Electronics Society's Games Innovations Conference 2009 (ICE-GIC 09)*, London, UK.
- [5] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, 586-591, 1991.
- [6] Batur, A.U., Flinchbaugh, B.E., & Hayes III, M.H., "A DSP-based approach for the implementation of face recognition algorithms." *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2, 253–256, (2003).
- [7] Wei, M., & Bigdeli, A. (2004). Implementation of a real-time automated face recognition system for portable devices. *Proceedings of the IEEE International Symposium on Communications and Information Technology*, 1, 89–92.
- [8] Moghaddam, B., & Pentland, A. (1997). Probabilistic visual learning for object representation. *Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 696–710.
- [9] Rowley, H.A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 23–38.
- [10] C. Gao and S. L. Lu, "Novel FPGA Based Haar Classifier Face Detection Algorithm Acceleration", *International Conference on Field Programmable Logic and Application*, pp. 373-378, Sept. 2008.

- [11] I. Sajid, M. M. Ahmed, I. Taj, M. Humayun and F. Hameed, "Design of High Performance FPGA Based Face Recognition System", PIERS Proceedings, pp. 504-510, Jul. 2006.
- [12] Y. Wei, X. Bing, C. Chareosnak, "FPGA Implementation of AdaBoost Algorithm for Detection of Face Biometrics", IEEE International Workshop on Biomedical Curcuits and Systems, pp. S1/6-17-20, Dec. 2004.
- [13] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models", Lecture Notes in Computer Science, vol. 1407, pp. 484-, 1998.
- [14] G. J. Edwards, T. F. Cootes, and C. J. Taylor, "Advances in active appearance models," in Proc. of the International Conference on Computer Vision (ICCV'99), 1999, pp. 137-142
- [15] T. F. Cootes, K. Walker, C. J. Taylor, "View-based active appearance models," Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on , pp. 227-232, 2000.
- [16] P. Huisman, R. van Munster, S. Moro-Ellenberger, R. Veldhuis, and A. Bazen, "Making 2d face recognition more robust using AAMs for pose compensation," Automatic Face and Gesture Recognition (FGR'06). 7th International Conference on, pp. 6 pp--, April 2006.
- [17] Y. Li, W. Ito, US Patent Application 20070071347.
- [18] H. Kameyama, W. Ito, US Patent Application 20060257047.
- [19] I. Bacivarov, M. Ionita, P. Corcoran, "Statistical models of appearance for eye tracking and eye-blink detection and measurement," IEEE Transactions on Consumer Electronics, August 2008.
- [20] I. Bacivarov, M.C. Ionita, and P. Corcoran, "A combined approach to feature extraction for mouth characterization and tracking", in Proc. of Irish Signals & Systems Conference (ISSC'08), Vol. 1, pp. 156-161, Galway, Ireland, 18-19 June, 2008.
- [21] I. Bacivarov, M. Ionita, and P. Corcoran, "Smart cameras: 2D affine models for determining subject facial expressions", submitted for publication to IEEE Transactions on Consumer Electronics, 2009.
- [22] I. Bacivarov, M. Ionita, P. Corcoran, "Statistical models of appearance for eye tracking and eye-blink detection and measurement," IEEE Transactions on Consumer Electronics, August 2008.
- [23] S. H. Seo, M. R. Azimi-Sadjadi, Bin Tian, "A Least-Squares Based 2-D Filtering Scheme for Stereo Image Compression", IEEE Transactions on Image Processing, vol. 9 issue 11, pp. 1967 - 1972, Nov 2000.
- [24] M. S. Moellenhoff, M. W. Maier, "DCT Transform Coding of Stereo Images for Multimedia Applications", IEEE Transactions on Industrial Electronics, vol. 45, issue 1, pp. 38 - 43, Feb 1998.
- [25] F. Davoine, M. Antonini, J. M. Chassery, M. Barlaud, "Fractal Image Compression Based on Delaunay Triangulation and Vector Quantization", IEEE Transactions on Image Processing, vol. 5, issue 2, pp. 338 - 346, Feb 1996.
- [26] J. D. Boissonat, O. D. Faugeras, E. Le Bras-Mehlman, "Representing Stereo Data with the Delaunay Triangulation ", Proceedings of IEEE International Conference on Robotics and Automation, issue 24-29, pp. 1798 - 1803, Apr 1988.
- [27] W. J. Kim, J. Kim, "Disparity Adaptive Filter for Anti-Aliasing of Stereoscopic 3D Images", 3DTV Conference, Transmission and Display of 3D Video, The True vision-capture, vol. 28, issue 30, pp. 41 - 44, May 2008.
- [28] S. Takezawa, G. Dissanayake, "Simultaneous Localization and Mapping Problems in Indoor Environments with Stereovision", 31st Annual Conference of IEEE, IECON 2005, Industrial Electronics Society, vol. 6, issue 10, Nov 2005
- [29] G. Roth, "Computing Camera Positions from a Multi-Camera Head", Proceedings of Third International Conference on 3-D Digital Imaging and Modeling, 2001, pp. 135 - 142, 2001.
- [30] A. Baumberg, "Reliable Feature Matching Across Widely Separated Views", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 774 - 781, 2000.
- [31] http://www.datasheetcatalog.com/datasheets_pdf/M/T/9/M/MT9M011.s.html
- [32] M. Romanca, "Embedded Systems", Lecture Notes, Transilvania University of Brasov, Romania, 2007
- [33] D. A. Patterson, J. L. Hennessy, "Computer Organization and Design, The Hardware/Software Interface", Fourth Edition.
- [34] R. Ramanath, W. E. Snyder, G. L. Bilbro, W. A. Sander III, "Demosaieking methods for Bayer Color Arrays", Journal of Electronic Imaging 11(3), pp. 306 - 315, July 2002.
- [35] D. Vanden Bout "VGA Signal Generation with the XS Board", application note, 25 Aug, 1998
- [36] Y. Fu, R. L. Huang and T. S. Danielsen, "Real-Time Avatar for Multimodal Human-Machine Interaction", IEEE Conference on Multimedia and Expo, pp. 991-994, Jul. 2007.
- [37] D. Kim and J. Sung, "Automated Face Analysis: Emerging Technologies and Research", Section 3.3 et. seq., Information Science Reference, Hershey, New York 2009.
- [38] Mircea Ionita, "Advances in the Design of Statistical Face Modelling Techniques for Face Recognition", National University of Ireland, Galway, 2008.
- [39] M. Feldman, "Wafer-Level Camera Technologies Shrink Camera Phone Handsets", Photonics Spectra. Vol. 41, no. 8, pp. 58-60. Aug. 2007
- [40] M. Ben-Ezra and S.K. Nayar,, "Motion Deblurring using Hybrid Imaging", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol.I, pp.657-664, Jun, 2003.
- [41] M. Ben-Ezra and S.K. Nayar, "Motion-based Motion Deblurring," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.26, No.6, pp.689-698, Jun, 2004.
- [42] Yu-Wing Tai, Hao Du, Michael S. Brown, Stephen Lin, "Image/video deblurring using a hybrid camera," cvpr, pp.1-8, 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008



Istvan Andorko received the B. Eng. Degree in Electronic Engineering from the "Transilvania" University of Brasov, Romania in 2008. He is currently pursuing a Ph.D. degree in Image Processing & Computer Vision at NUI, Galway. His research interests include image and signal processing, VLSI and embedded systems..



Petronel Bigioi received B.S. and M.S. in Electronic Engineering in 1997, 1998 and Ph.D. in Digital Still Camera Connectivity in 2005 all from the "Transilvania" University of Brasov, Romania. He received M. Eng. Sc.. degree at National University of Ireland, Galway in 2000. Currently he is VP of Engineering at Tessera Ireland Ltd. His research interests include VLSI design, digital imaging, communication network protocols and embedded systems. He is a member of IEEE.



Peter Corcoran received the BAI (Electronic Engineering) and BA (Math's) degrees from Trinity College Dublin in 1984. He continued his studies at TCD and was awarded a Ph.D. for research work in the theory of Dielectric Liquids. He is currently Vice-Dean of research in the College of Engineering & Informatics, National University of Ireland Galway. His research interests include embedded systems, home networking, digital imaging and wireless networking technologies.