SwissRanger SR3000 and First Experiences based on Miniaturized 3D-TOF Cameras

Thierry Oggier*, Bernhard Büttgen, Felix Lustenberger
Swiss Center for Electronics and Microtechnology, CSEM, Badenerstrasse 569, 8048 Zurich, SWITZERLAND

Guido Becker
IEE, Zone industrielle Findel, 2b, route des Trêves, 2632 Luxembourg, Luxembourg

Björn Rüegg, Agathe Hodac
Fachhochschule Rapperswil, Oberseestrasse 10, 8640 Rapperswil, SWITZERLAND

Abstract
The latest development in range imaging 3D-cameras is presented, the so-called SwissRanger3000 (SR3000). The SR3000 includes a modularly built electronics stack. This modular stack allows for easy adjustments to the camera for specific customer requirements. The core of the SR3000 is comprised of a background suppressing 3D-sensor with QCIF resolution. Furthermore, the first evaluation results of field-tests using a 3D-camera developed at CSEM are presented. The outcome of the evaluation phase for the automotive industry is discussed. A new development using the camera as detecting device for a virtual interaction on a large screen has successfully been implemented. Finally, the latest breakthrough in distance accuracy enables the 3D-TOF camera to enter the biometrics markets.

Keywords
SwissRanger, 3D camera, time-of-flight principle, range-imaging camera, 3D sensor

1 INTRODUCTION

During the last few years, the first all-solid-state 3D-cameras based on the time-of-flight principle became available on the market [1, 2]. The cameras are at affordable prices targeting feasibility tests for industrial interests. For the first time, customers and potential users of the 3D-camera could purchase such units and evaluate them for their purposes and applications. New problems have been solved and more cost-efficient solutions have been implemented. The latest release of a 3D-camera, the SwissRanger3000 (SR3000), is presented in section 2. The camera available as an evaluation kit also fulfils the requirements for integration into larger devices to be manufactured in medium volume application. The characteristics and setup of the SR3000 camera are presented within the framework of this paper.
In addition to the presentation of the state-of-the-art 3D-camera, various experiences gained with the proceeding versions of SR3000 are discussed in sections 3, 4 and 5. First applications realized by industrial and universal collaborators are presented.
The automotive sector is one of the industries leading the way in the development of 3D-cameras. Their targeted applications of the 3D-camera such as the out-of-position detection

(OoP) and passenger classification for controlled airbag deployment as well as pedestrian protection are discussed. The state-of-the-art 3D-camera for the automotive interior sensing developed in collaboration by IEE and CSEM is presented in more details.

A second application implemented in collaboration between the University of Applied Sciences in Rapperswil, Switzerland and CSEM aims for proofing the feasibility of a new interactive input device. In the world of virtual interaction (human-machine interface), a new virtual touch screen for very large screens is demonstrated. This development is foreseen to find use during exhibitions where the visitor can interactively control the screen, finding information about a company, or even play games.

Based on the latest progresses in 3D-imaging, completely new applications can now be targeted with low-cost time-of-flight 3D-cameras. In section 5 an example of a new field of application is described: biometrics.

The paper concludes with an outlook in the world of 3D-imaging at CSEM.

2 SWISSRANGER3000

The SwissRanger3000 camera is based on a phase-measuring time-of-flight (TOF) principle. A light source emits a near-infrared (NIR) wave front that is intensity-modulated with a few tens of MHz. The light is reflected by the scene and imaged by an optical lens onto the dedicated 3D-sensor. Depending on the distance of the target, the captured imaged is delayed in phase compared to the originally emitted light wave. Measuring the phase delay, the distances of the complete scene can be determined. The result of the acquisition is a depth map of the scene. The principle is described in more details in [3].

The SwissRanger3000 is based on the latest technologies implemented for time-of-flight ranging devices. The core of the camera is the integrated sensor. The sensor has been manufactured by ZMD [4]. The manufacturing process has been optimized regarding demodulation speed as well as noise performance. Based on this manufacturing technology a sensor is obtained that is specifically optimized for range imaging. The camera electronics controls the sensor, processes the images and transfers depth maps to the computer. It has been set up as a stack of three printed-circuit-boards (PCBs). The illumination is based on an array of light emitting diodes (LED). LEDs were chosen targeting very low system cost at high volumes.

The optical lenses and the filter are not custom-designed but were carefully selected to fulfill their specifications the best possible.

2.1   Integrated 3D-TOF image sensor

The 3D-TOF sensor has been designed in an optimized manufacturing process at ZMD. In close collaboration, ZMD and CSEM adjusted and modified the technology process targeting a highly sensitive pixel providing demodulation speed of a few tens of MHz to about 100 MHz. More about the modifications in the technology and the technology itself can be found in [5]. The recently implemented sensor exhibits QCIF resolution (176 x 144 pixels) with each pixel characterized by high quantum efficiency together with extremely well-controlled demodulation performance. The pixels allow an almost noise-free accumulation of charges over thousands of periods and for subtracting the electrons generated by the background light. This pixel-inherent background suppression enables for the first time the use of the 3D-TOF camera in outdoor applications.

In the camera manufacturing process, the sensor can completely be tested before packaging. This allows significant reductions in costs and time during the camera production.

Self-test features are implemented on the sensor to verify each single transistor down to the pixel demodulation functionality.

2.2   3D-camera electronics

The electronics of the camera is modularly set up in a stack of three printed circuit boards. The top PCB hosts the sensor itself, and therefore is called "sensorhead". It controls the sensor and supplies it with the power. The different bias voltages and currents used for the sensor operability are implemented on the "sensor head" board as well as the signal conditioning for the data transfer to the next PCB. This includes the A/D conversion of the sensor outputs.

The data acquired by the "sensor head" is transferred to the IPM (image processing module) located underneath the sensorhead. The IPM processes all the images and calculates the distances out of the raw data delivered by the "sensor head". It determines the qualities of the acquired distances in real-time and can apply first image filters and/or execute application-dedicated algorithms. The processed data is then transferred to the "communication interface module" (CIM).

The CIM contains the interface between the camera and the computer. As a first implementation, the data transfer link is based on USB2.0. Thanks to the modular set-up of the camera, the CIM can easily be modified to support customized communication protocols. Other commonly-used interfaces such as Ethernet and CAN-Bus are in development.

Figure 1a shows the photograph of the PCB stack with mounted IPM, CIM and sensorhead module.
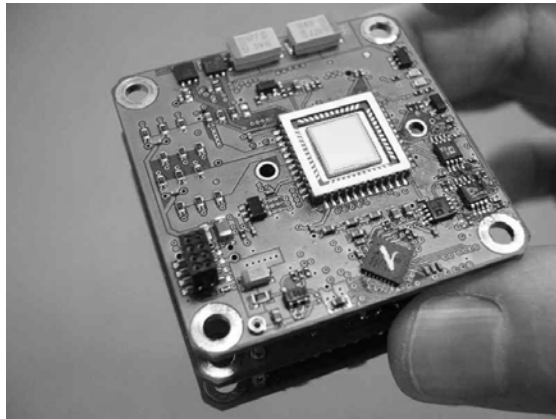
2.3   3D-camera illumination

The illumination for the SR3000 has been chosen based on the customer criteria of developing a cost-efficient and eye-safe solution. Therefore, more expensive implementations based on laser, laser diodes or vertical cavity surface emitting lasers were not explored. The illumination unit of the SR3000 camera consists of an array of 55 LEDs. The LEDs are carefully selected to fit the requirements of the 3D-camera.

The illumination unit is driven at a few tens of MHz. At the same time, the light power has to be as high as possible, but still lying inside the eye-safety norms. Additionally to these requirements, the modules have to be small, efficient and cover as uniformly as possible the same field-of-view as it is captured by the camera.

The final manufactured illumination unit uses LED with a central wavelength of 850 nm and a total mean emitted power of 1 W.

Figure 1b shows a photograph of the LED illumination unit.
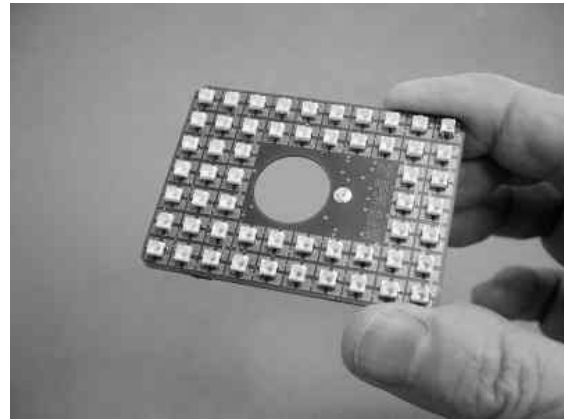
a)   **b)**

Figure 1: Photograph of a) the PCB stack as implemented in the SR3000 camera and b) the illumination unit board delivering a total mean light power of 1 W.

2.4   3D-TOF camera characteristics

The SR3000 resolves distances down to a few millimeters. Due to the almost noise-free demodulation capability of the sensor's pixels, the distance resolution virtually approaches the physical limitations given by the photon shot noise.

The user's control of the camera is enabled via the USB2.0 interface. The main parameters of the camera such as integration time can be set and read out. By setting the parameters, the operator can choose between different acquisition modes: one for high speed 3D-image acquisition reaching a frame rate of about 40 frames per second at distances of a few meters, or one for more precise image acquisitions that still achieves about 20 frames per second for scenes in a range of a few meters.

The non-ambiguous range determined by the modulation frequency can be extended to more than 30 meters by appropriately setting the camera registers. This allows the implementation of the SwissRanger3000 camera for short range applications as well as for longer range applications.

The camera is temperature calibrated and can be self-tested inside the manufacturing process. Firmware updates can directly be downloaded from the SwissRanger website [2] through the USB port. The camera has an overall size of 50 mm x 48 mm x 67 mm.

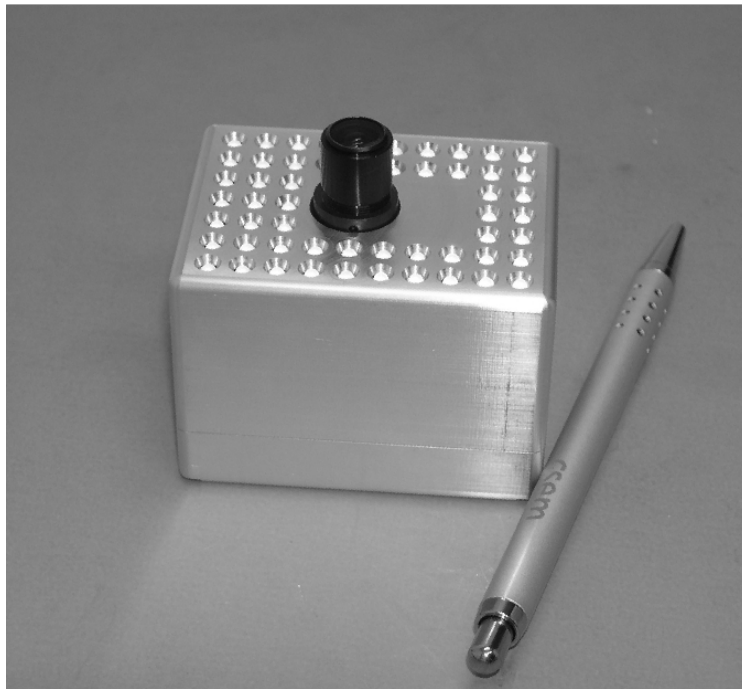The picture in Figure 2 depicts the 3D-camera SR3000.

Figure 2: Photograph of the SwissRanger3000 camera.

3 AUTOMOTIVE APPLICATIONS

3.1 Passive safety Out-of-position Detection

The first application of the 3D-technology under development is an occupancy monitoring system. The goal of this system is to fulfill the requirements of the US Federal Motor Vehicle safety requirements (FMVSS 208). These requirements describe demands towards a static sensing system and include an option for a dynamic sensing system. The first law for a static system is active since September 2003.

The US law requires for the first phase from September 2003 on smart airbag systems. Occupant sensing systems must guarantee an airbag deployment in presence of an adult. For children of different age classes (1 year, 3 year, 6 year), with and without child-seats, airbags must be either suppressed or must deploy according the OEM's airbag deployment strategy. The occupant sensing systems have to provide the corresponding information. The FMVSS 208 includes an option called DASS (dynamic automatic suppression system), targeting an out-of-position (OOP) detection of the occupant. OOP aims at defining the seating position and posture of the passenger, to obtain their relevant distance to the airbag unit.
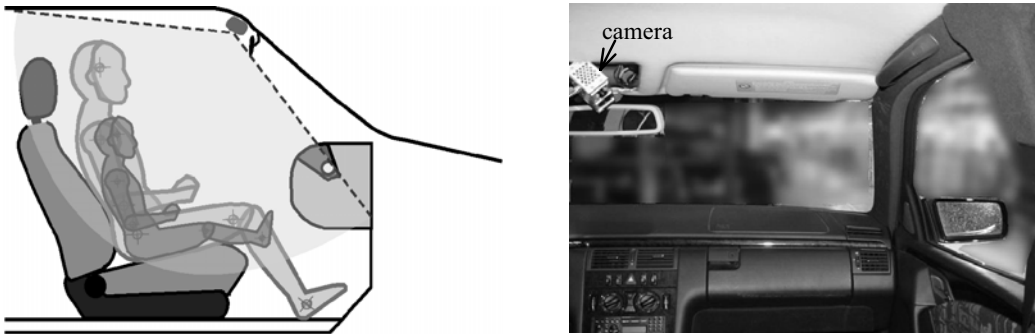
Figure 3: The occupant sensing system installed in a vehicle: On the left the field of view is sketched. The photo on the right shows the installation of a prototype in a test vehicle near the review-mirror.
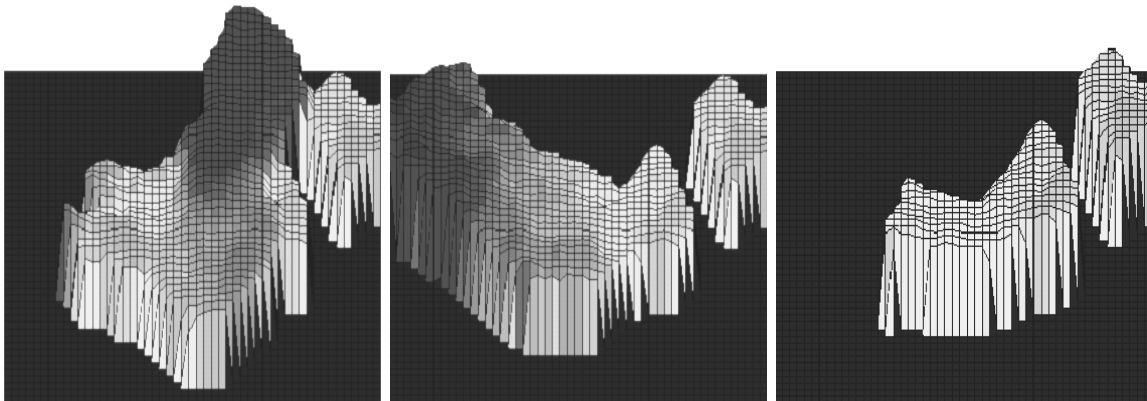

Figure 4: Sample distance images of the 52x50 pixel camera, with background compensation applied. The closer the distance of a point on the object to the camera the more it stands out from the uniform background. From left to right: adult (facing left), rear-faced infant seat, empty seat. In all three images the headrest is clearly visible on the right.

The advantages of the 3D sensing can be seen in the Figure 9, with a clipping operation eliminating the background in the picture, as e.g. doorframe or A-/B-pillar. Only the objects in the relevant distance range are considered. The closer the distance of a point in the scene to the camera, the more the related distance pixel stands out from the uniform background. This clipping method allows reducing the number of data to be considered for further processing, by discarding regions in the field of view belonging to the background.

For occupant classification, features need to be extracted that contain the crucial information in each frame.

For this purpose we use descriptors that are either derived from the range frame itself or from the representation of the data in the Cartesian vehicle coordinate system.

Shape features can be calculated directly from a binarized 2D range image, e.g. an image where all background pixel are set to zero while all non-background pixels to one. By keeping only pixels in the vicinity of a discontinuity in range we can calculate an edge image, for which contour descriptors can be derived, e.g. area, height and orientation of ellipsoidal contours. Further features can be gained from the distribution of scatter points in the 3D vehicle coordinate system. One can then fit shapes, (e.g. of ellipsoids) to clusters of points, which represent the object. In total around 15 features are extracted, which then form the input of the classification algorithm.

The goal of the classifier is to discriminate forward facing (FF) or rear facing (RF) child restraint systems, adults (P) and empty seats (E). To study the suitability of the 3D-camera technology for this classification, in total 150 sequences were recorded for several adults (weights from 50 to 90 kg) in different postures and a dozen different child seats.

3.2 Feasibility of Exterior Applications

First examinations for sensing systems for exterior applications were made with a 3D camera mounted on the car bonnet. With this setup sequences for near-field applications were recorded. A system simulation to analyze the technical requirements for exterior applications is currently in realization.
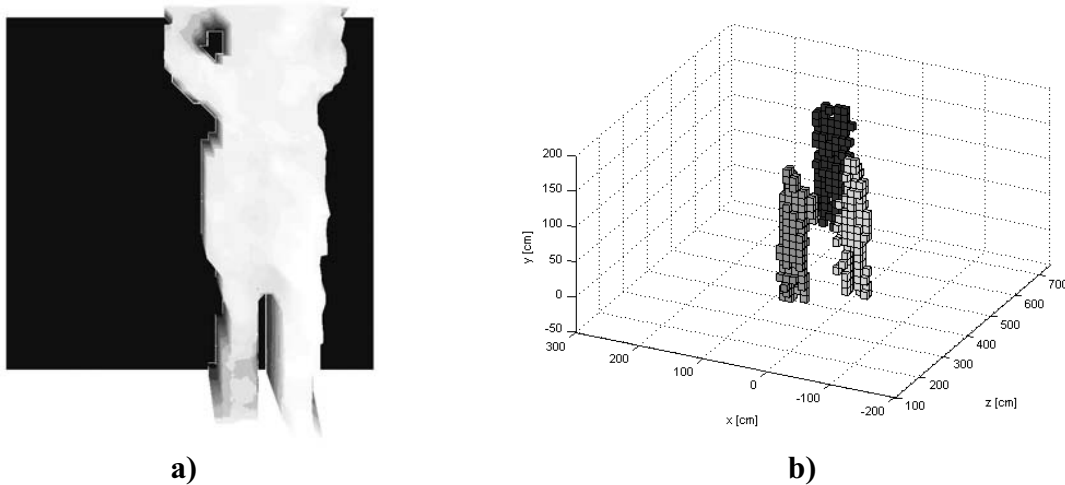

a)   b)

Figure 5: Outside 3D picture 50*52 pixels e.g. for pedestrian protection.

The analysis of possible applications has started. The actual system has a reduced performance for such exterior applications due to the fact that it is designed for an interior application with distance ranges below 2 m. In order to realize an improved illumination the use of Vertical Cavity Surface Emitting Laser (VCSEL) elements or other laser elements is targeted. With this improvement it will be possible to achieve a distance measurement up to the non-ambiguity range of 7.5 m for the 20 MHz modulation. With a reduced modulation frequency the non-ambiguity range can be extended to 15 m and higher.

3.3 Automotive Exterior applications

Besides the occupant monitoring functionality, there is a variety of application in the car exterior that can be addressed with the 3D-camera technology. Every application will have its own requirement regarding the FOV, the range and the lateral sensor resolution. These requirements need to be evaluated. In general, the demand for a larger range requires a more powerful illumination than for the interior application.
With larger distance ranges, the applications around the car sketched in Figure 8 can be addressed with the 3D sensor technology.
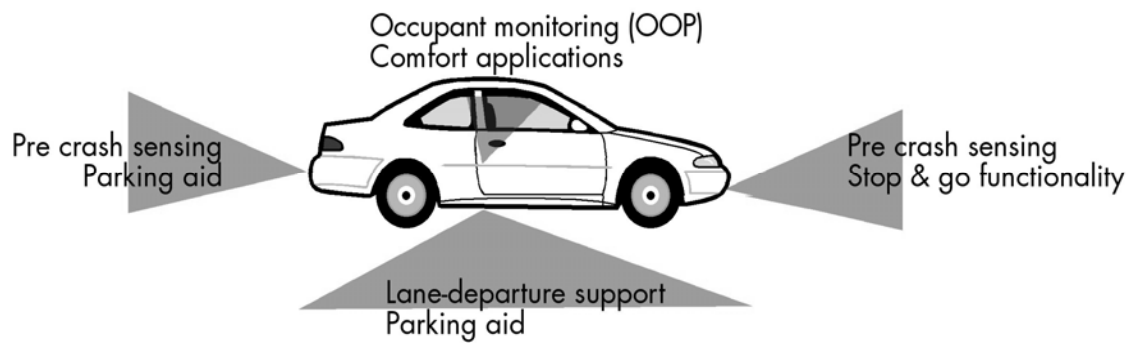
Figure 6: Exterior applications for the 3D TOF sensor

The illustration above shows a mix of different comfort and safety applications. The main focus for exterior systems is set to the applications with the highest benefit for the car owner. In this field all the developments aim at the near range up to 15 m around the car.

3.4 Automotive Outlook

The actual implementation of the 3D TOF sensor shows a high distance measurement quality. The basis of the technology is presented. Serial applications can now be realized. The first application IEE is working on is the occupancy monitoring in the field of passive automotive safety. In the near future this application will help to improve the safety in the car by fulfilling the legal requirements of the American FMVSS 208. Applications for active safety, as e.g. stop & go functionality and pre-crash sensing in the near field of up to 15 m around the car, are a field of high interest for the optical 3D TOF technology.
The use of a more efficient semi-conductor technology will allow the integration of more functionality into each pixel or/and an increase of the sensor resolution. Improvements regarding the optical fill-factor of the pixel structure and the use of micro lenses will result in improvements of the optical power budget and/or the sensor size.
The development of algorithm structures and principles for reliable processing of the 3D data of a scene is one of the main topics of the system development.
Algorithms on the basis of a 3D representation of a scene allow complete new approaches. Compared to 2D interpretations, the third dimension will increase the robustness of object classifications. The availability of distance information allows using a lower resolution sensor type.
The image processing development includes the analysis of different algorithm principles. Structures as neuronal networks, linear classifiers and shape based realizations for different applications are under development.

4 LARGE SCALE VIRTUAL INTERACTIVE SCREEN

Commonly known touch-screens use tactile sensors located behind the screen to detect the human interaction. These sensors limit the screen size. A new approach using an adjusted version of the SwissRanger 3D-camera allows human-machine interaction on large scale screens. Contrary to conventional touch-screens, tactile sensors located behind the screen are no longer required.
Such large scale interactive "touch" screens could be used at exhibitions, as marketing instruments or as customer information screens in shops. The development can even be further refined to open interactive gaming applications.

4.1 Interactive screen setup

In the first demonstrated use of the virtual interaction screen, a beamer projects the computer screen on a wall. The projected screen has a size of about 1 m x 1.5 m. The camera is mounted on the left corner of the screen and captures the scene in front of the screen.
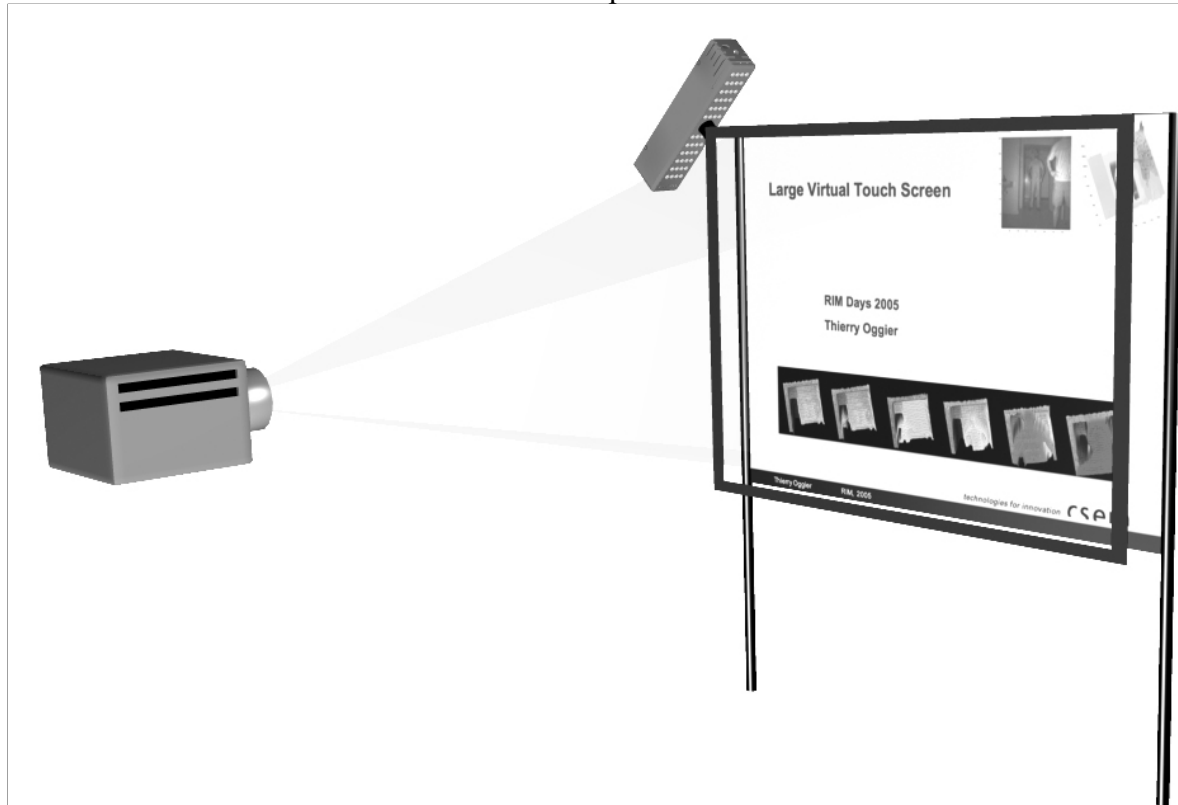

Figure 7: Setup of the virtual interaction screen.

The task of the SwissRanger 3D-TOF camera is to detect and locate the movements of the user's hand touching the screen. Based on the movements, control sequences (e.g. to control a power point presentation) are sent to the computer.

4.2 SwissRanger implementation

First of all, the SwissRanger camera has to be calibrated. The calibration is performed following the procedure proposed in [6]. The camera has been modified in order to illuminate and capture just ten rows of pixels. The limited number of lines is absolutely sufficient to provide high reliability and to detect the motion directions. Each row consists of 160 pixels. If the emitted light is focused on ten lines and only these ten lines of the sensor are readout and transferred to the computer then the amount of data to be transferred has been significantly reduced and the overall frame rate significantly increased. Frame rates up to 50 Hz have been achieved.

An algorithm to automatically select the best integration times for the image acquisition has been developed. The algorithm takes into account the reflectivity as well as the distance of the object and consequently deduces the most appropriate exposure time required to obtain the best distance accuracies.

4.3 Virtual interaction screen functionality

The virtual interaction screen, as implemented in a first version, controls a PowerPoint presentation. Algorithms for the detection of the mouse functions such as "forward", "backward" or "virtual click" on a button have been developed.

A second implementation of the virtual interaction screen targets gaming application. The widely known game "snake" is controlled by the operator in front of the screen.
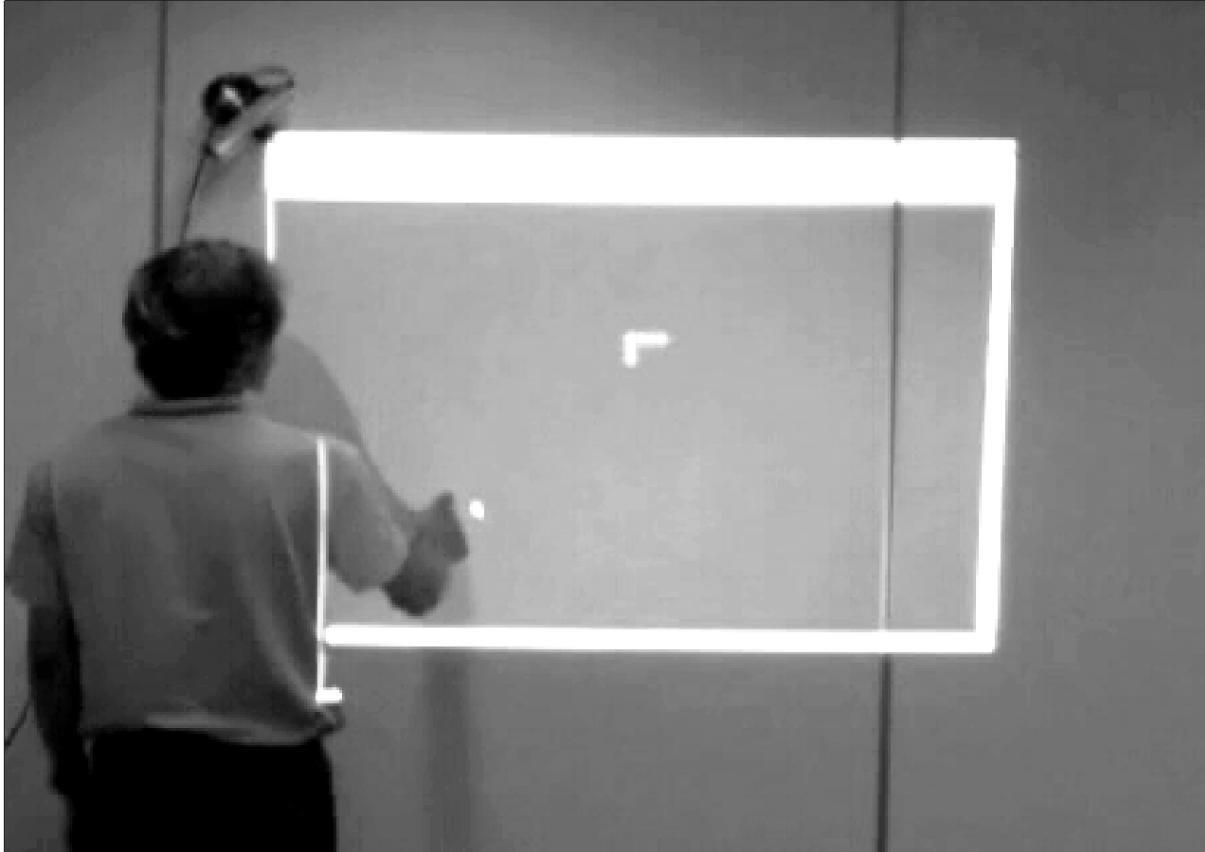

Figure 8: Virtual interaction screen with a player controlling the "snake" movements.

5 BIOMETRICAL APPLICATIONS

Recent developments and tests have shown that the acquisition of the third dimension in combination with the intensity image significantly increases the face recognition robustness of people identification devices [7]. The so-called multi-modal face recognition approach (2D and 3D) is based on the acquisition of two separate imaging devices: one capturing a black and white image and the second is used for the three dimensional image. This multimodal approach is commonly reported to outperform a single modality. So far, due to the high resolution requirements, the 3D-face acquisitions were performed by expensive and mostly bulky laser scanning devices. The biometrical industry will benefit from a cost-efficient, robust system that delivers 2D as well as precise 3D information.

Recently, the improvements in distance resolution achieved and demonstrated with the 3D-TOF camera developed at CSEM enable the entry of the biometrics market with a low-cost, high-speed, miniaturized and robust solution for the acquisition of 2D and 3D images. Implementing the appropriate temporal filtering, distance resolutions as low as one millimeter can be attained.

Two examples of high-resolution 3D image acquisitions are illustrated in Figure 9. As a first example, a foot has been captured with the SwissRanger 3D-camera. The acquisition for the image requires a few 100's of milliseconds. On the left side of Figure 9, the depth image of the foot acquired by the SwissRanger camera is illustrated. The depth is coded in a black-and-white scheme, where black regions correspond to smaller distances than the brighter ones. The quality of the distance map already fulfills the specifications required for orthopedic applications. Regarding the face recognition system, a second image is illustrated in Figure 9b. The depth map of the acquired face is again coded in intensity and quantified by the colorbar.
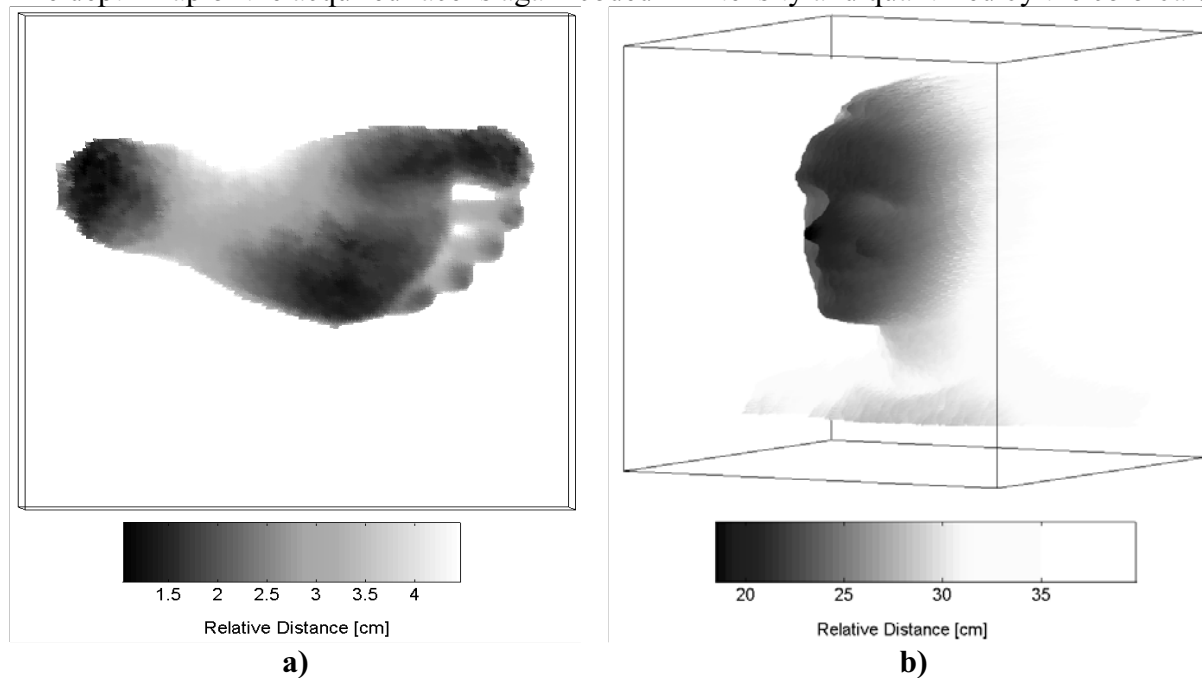

a)   b)

Figure 9: First biometrical experiences using 3D-TOF acquisition of a) a foot and b) a face.

6 CONCLUSIONS

In this paper, the newly miniaturized 3D-camera SwissRanger3000 as well as first experiences with 3D-camera prototypes have been presented. The SR3000 camera acquires in real-time depth maps of the scene with a QCIF sensor resolution. It can operate in strong background light conditions up to bright sunlight.

Experiences gained with the first 3D-TOF camera prototypes available on the market are described. The first mentioned application derives from the automotive industry. The 3D-camera might find several uses in the automobile. The first targeted application is the out-of position detection of the passenger and/or driver for a better airbag deployment control. The proof of principle has already been concluded for the DASS (dynamic automatic suppression system) and the industrialization has been started. Other applications have followed e.g. pedestrian protection and parking aid. Another development presented is the implementation of the 3D-camera SwissRanger as virtual interaction screen. Thanks to the 3D-camera, large screens can interactively be controlled. The interactive operation of a PowerPoint presentation as well as a first simple game are successfully implemented and presented.

The latest results in 3D-TOF imaging show that the market of biometrics can be envisaged with TOF cameras. Examples with millimeter resolution targeting face recognition and orthopedic application are illustrated.

REFERENCES

1. T. Oggier, M. Lehmann, R. Kaufmannn, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, N. Blanc, An all-solid-state optical range camera for 3D-real-time imaging with sub-centimeter depth-resolution (SwissRanger), Proc. SPIE Vol. 5249, pp. 634-545, 2003.
2. http:\\www.swissranger.ch
3. R. Lange and P. Seitz, "Solid-State Time-of-Flight Range Camera", IEEE J. Quantum Electronics, Vol. 37 (3), 390-397, March 2001.
4. http:\\www.zmd.de
5. T. Oggier, R. Kaufmann, M. Lehmann, P. Metzler, G. Lang, M. Schweizer, M. Richter, B. Büttgen, N. Blanc, K. Griesbach, B Uhlmann, K.-H. Stegemann, C. Ellmers, "3D-Imaging in Real-Time with Miniaturized Optical Range Camera", Opto Conference Nurnberg, 2004.
6. T. Kahlmann, H. Ingensand, "Calibration and improvements of the high resolution range-imaging camera SwissRanger[TM]", Proceedings of the SPIE, Vol. 5665 No. 20, 2005.
7. K. Chang, K. Bowyer, and P. Flynn. "An evaluation of multi–modal 2D+3D face biometrics" IEEE Trans. on Pattern Analysis and Machine Intelligence, 27(4):619–624, Apr. 2005.