# Université Laval Face Motion and Time-Lapse Video Database (UL-FMTV)

by Reza Shoja Ghiass*, Hakim Bendada*, Xavier Maldague*

*Computer Vision and Systems Laboratory, Laval University, Quebec City (Quebec) G1V 0A6, Canada, reza.shoja-ghiass.1@ulaval.ca, bendada@gel.ulaval.ca, maldagx@gel.ulaval.ca

**Abstract**

In this paper, we present Université Laval Face Motion and Time-Lapse Video Database (UL-FMTV), the largest facial video database of the world in the Mid-Wave Infrared spectrum. The lack of a large video database gathered during multiple sessions within a relatively large period which contains images of subjects with different ethnicities, ages and sexes are missing in the literature. In this database, we not only focused on these factors but also considered a wide range of facial poses and expressions. The database is obtained from 238 subjects and is available for public use for research purposes only.

## 1. Introduction

In the course of the last three decades, visible face recognition has been one of the most interesting research topics in the domain of computer vision, pattern recognition and most recently, deep learning. However, the success of face recognition was never without challenge. Visible face recognition system has reached a great level of maturity with some practical success. However, a range of challenging factors has posed serious problems to the visible face recognition systems. Most notably, illumination, facial pose, and facial expression are among the most challenging factors.

A large number of previous works have focused on finding solutions to overcome the aforementioned challenges. For example, the facial pose has been normalized by fitting a facial model (i.e., a set of fiducial points) using synthetic approaches or statistical methods [1-3], while illumination normalization has been accomplished using image processing filters [4, 5] and statistical facial models. Using other modalities than visible spectrum such as 3D imaging or thermal imaging are another solution to cope the illumination dependency problem. 3D data obtained from 3D scanners are less dependent on the illumination changes. They can also handle the rotated faces in multi-view applications, but they have some disadvantages. The cost of such systems is high, their processing speed is low and some artifacts are produced due to speculation [6]. Thermal infrared sensors can measure the emitted heat energy and they are not dependent on lighting conditions. However, thermal imaging for face recognition suffers from some problems. It is subject to environmental temperature, emotional, physical and health conditions. Another problem of the thermal spectrum is the opaqueness to eyeglasses. This makes a large portion of the face wearing eyeglasses to be occluded in thermal images [7]. Consequently, some information around the eyes will be loosed which can decrease the accuracy of the system. To cope all these problems together, a large number of publications in the literature have borrowed the same methods in the visible face recognition and used them in a fusion based framework with thermal modality. Only a few numbers of IR-specific works have been reported in the literature [8-12].

Much like visible spectrum, both pose and facial expression changes present a major challenge to the current state of the art. In the visible spectrum, one of the most promising fields to normalize both factors is fitting some generic models to the test images. A very successful example of such models in the visible face recognition domain is called Active Appearance Models (AAMs) [13] or the revisited version of them called Inverse Compositional AAMs [14]. However, such models need to be trained first to capture both facial pose and facial expression and express them in terms of some principle components as well as a mean vector for both shape and texture. To reach this goal BASEL database [15] for example uses faces of 100 male and 100 females to train 3D generic face models. Next, this generic model can be used to be fitted to the face of any test subject.

Unlike visible modality, training some generic face models in the thermal domain had not been accomplished until our previous works. The main reason for this problem was the lack of large databases which not only contains the face of subjects from different ages, races, and sexes but also include as much as frames possible so that one can capture facial deformations when facial pose or facial expression changes. For this purpose, we gathered the first largest thermal video database of the world, Université Laval Face Motion and Time-Lapse Video Database (UL-FMTV). We also did training and testing of IC-AAMs on this database for the first time in the world (Fig. 1). This database contains facial videos with high resolution in the MWIR sub-band which makes it unique in the world from this standpoint too.

Figure 1: An AAM fitted to a facial image from our database

We believe the UL-FMTV database is an important contribution to the face recognition community, and we, therefore, encourage researchers to develop thermal face recognition algorithms and to report results using this data. This document is organized as follows: Section 2 briefly describes the existing facial databases in the thermal modality. The recording methodology and the different recording sessions included in the dataset are explained in section 3. Database specifications is explained in section 4. Finally, we draw the conclusions in section 5.

## 2.   Infrared Face Databases

In this section, we review some of the most important infrared databases Table 1 summarizes the most important thermal infrared face databases.

**Table 1.** *A summary of the main databases which contain facial images acquired in the thermal infrared spectrum [16].*

| Name | Subjects | Variability | | | |
|---|---|---|---|---|---|
| | | Illumination | Pose | Expression | Time-lapse |
| Equinox | 90 | ● | ○ | ● | ○ |
| IRIS | 32 | ● | ● | ● | ○ |
| IRIS-M3 | 82 | ● | ○ | ○ | ◑ |
| UND | 241 | ◑ | ○ | ◑ | ● |
| UH | 138 | ○ | ● | ● | ○ |
| SCFace | 130 | ● | ● | ◑ | ◑ |
| FSU | 10 | ○ | ● | ● | ○ |
| UC | 200 | ○ | ● | ◑ | ● |

### 2.1   Equinox

The "Human Identification at a Distance" database, collected by Equinox Corporation has been the most used data set for the evaluation of infrared-based face recognition algorithms in the literature. The data set contains 240x320 pixel images of 90 subjects in the visible, LWIR, MWIR, and SWIR sub-bands. Note that all images of each subject were acquired in a single session making it unsuitable for the evaluation of robustness to time-lapse.

### 2.2 IRIS Thermal/Visible

IRIS Thermal/Visible database is a database of both visual and thermal images. It is collected across pose, illumination and expression variation. The set comprises 4228 pairs of 320x240 pixel images. There are 32 individuals in the database. Much like the Equinox database, all images of each subject were acquired in a single session.

**2.3 IRIS-M3**

The IRIS-M3 is a face database which contains both thermal and visible spectrum images. This database also includes multispectral images acquired in 25 sub-bands of the visible spectrum. The database contains images of 82 individuals of various ethnicity, age and sex, and a total of 2624 images with 640x480 pixel resolution. This database was acquired in two sessions. The IRIS-M3 database does not contain any pose or expression variation.

**2.4 University of Notre Dame (UND)**

The University of Notre Dame database contains LWIR and visible spectrum facial images with 320x240 pixel resolution of 241 individuals under two illumination conditions. The database was collected in multiple sessions and it contains a total of 2492 images.

**2.5 University of Houston (UH)**

The University of Houston database consists of a total of 7590 thermal images of 138 subjects, with a uniform distribution of 55 images per subject. Subjects are of various ethnicity, age, and sex.

**2.6 Florida State University (FSU)**

The Florida State University comprises 234 images with 320x240 pixel resolution of 10 different subjects across a range of poses and facial expressions.

**2.7 UC Irvine Hyperspectral (UC)**

Hyperspectral images of 200 subjects exist in the University of California/Irvine database. All images were captured with 468x494 resolution. Subjects were imaged with a neutral facial expression in the frontal, profile and semi-profile poses, as well as with a smiling expression in the frontal pose only.

**3.    Data Acquisition**

In this section we first describe our recording methodology and then describe the different recording sessions constituting the dataset.

**3.1.  Recording Setup**

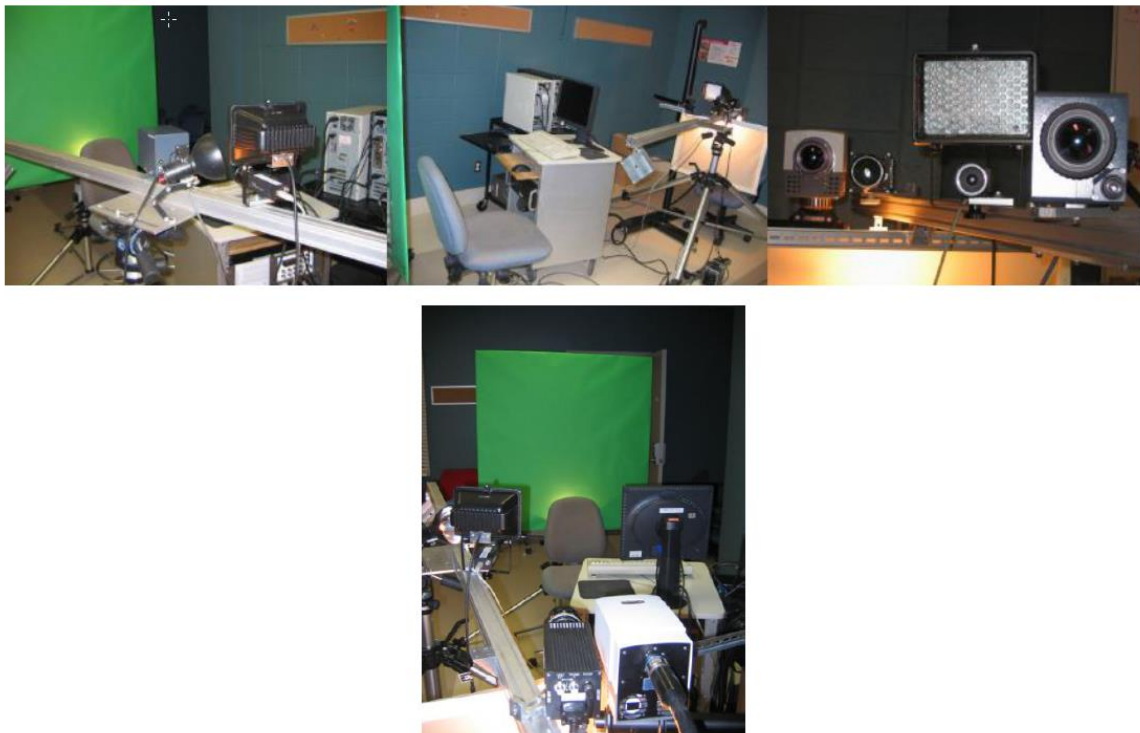The recording setup is as shown in Fig. 2. It comprises the following cameras:



Figure 2: The imaging setup used to gather the UL-FMTV face database
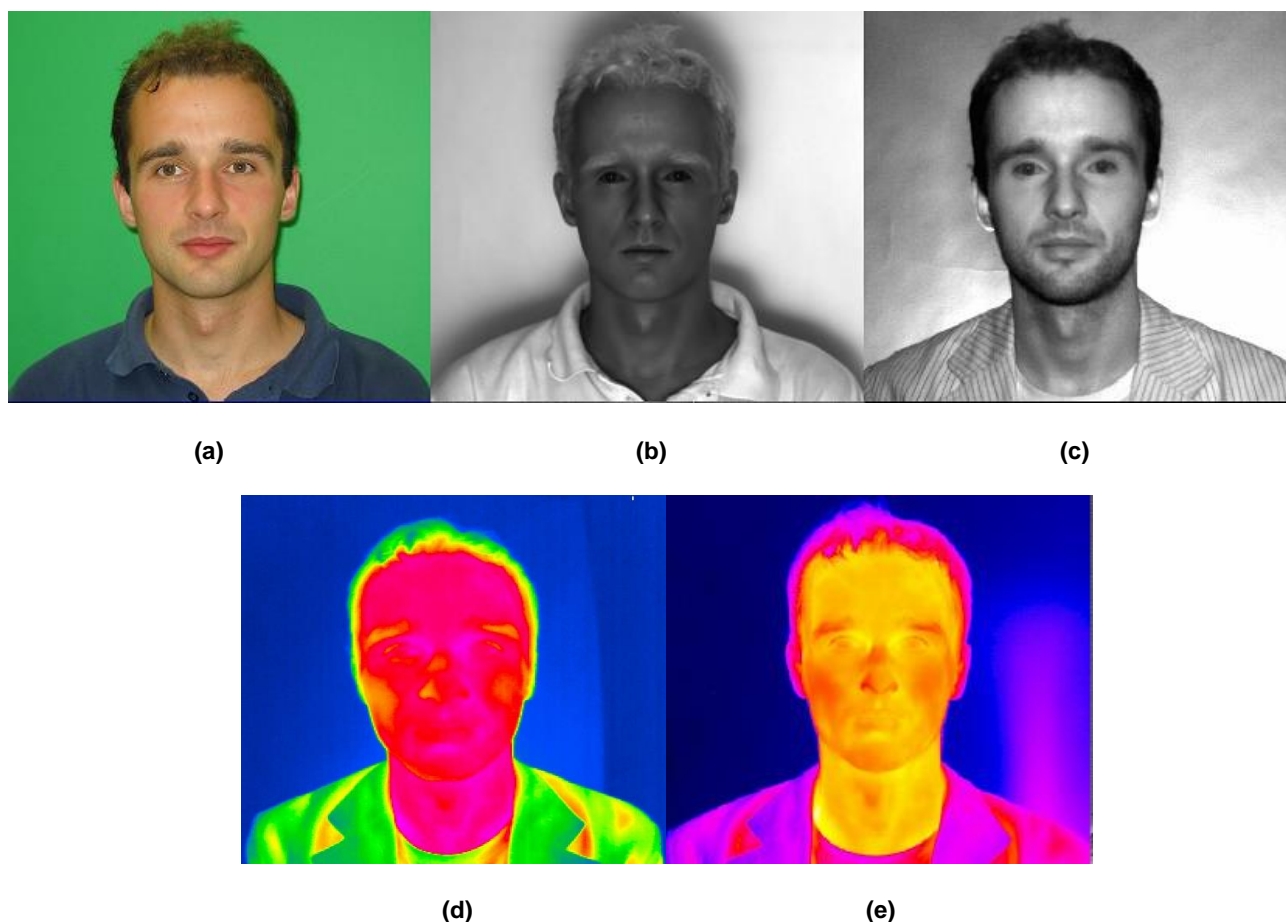
Figure 3: A subject from UL-FMTV: a) Visible, b) SWIR, c) NIR, d) LWIR and e) MWIR

- LWIR thermal camera: This camera is manufactured by Jenoptik and it operates at 8000-14000 nm.

- MWIR thermal camera: This camera is a Phoenix Indigo IR camera produced by FLIR. It is a cooled, Indium Antimonide InSb focal plane array camera, with Noise Equivalent Temperature Difference (NETD) of 0.025K at room temperature. The camera operates at 3000-5000 nm.

- SWIR camera: It is a scientific CMOS camera which is made by Goodrich. This camera operates in the range of (900-1700 nm).

- NIR/Visible camera: This camera is a standard CCD one made by Mutech, and it operates at 750-1100 nm.

Figure 3 demonstrates a subject acquired using this imaging setup. Note that currently, only the MWIR images of our database is available for the public.

## 3.2 Acquisition Sessions

In order to evaluate different aspects of an infrared face recognition algorithms, we designed a set of recording sessions, each one characterized by a combination of four main variables that can affect the accuracy and that we describe below: time-lapse, head pose, facial expression, eye-glasses, temperature. This database has been gathered in several sessions between 2010 and 2014. Figs. 4-7 demonstrate some samples from different sessions in the database.

## 4 Database Specifications

In this section we briefly explain some details about the database.

**4.1 Database Separation**

There exist two folders in the database. One folder contains the video sequence of subjects which will be used to train and test face recognition algorithms. To be consistent with the literature [17], we named this folder Genuine. The Genuine folder consists of 134 subjects: 86 males and 48 females. In this database, 44 subjects out of the 134 subjects participated in multiple sessions ranging from several weeks to 4 years. Among these 134 subjects, 44 subjects have been appeared with and without eyeglasses. All the subjects were required to change the head pose from left to right and show some facial expressions when videos were captured. This folder will be used to calculate the CMC curves.

The other folder is intended to challenge face recognition system in a verification scenario, i.e., calculation of ROC curves. We named this folder impostor. This folder consists of 104 subjects. The reader is referred to Appendix E in [17] to be familiarized with these two terms. Next, the performance metrics of a face recognition system, i.e., ROC and CMC curves can be calculated based on these terms.
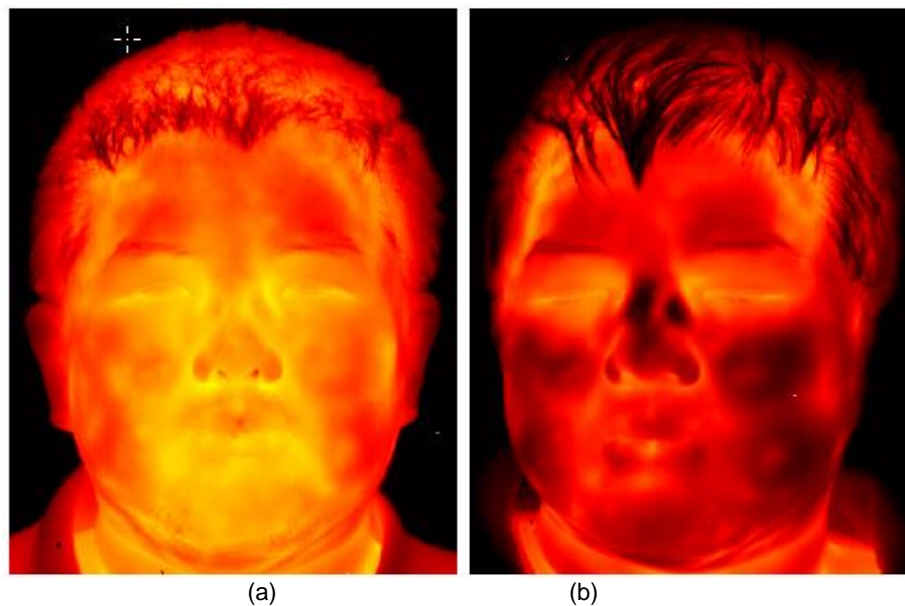


(a)                                  (b)

Figure 4:   A sample of UL-FMTV after time-lapse: (a) Room temperature, and (b) after exposure to cold
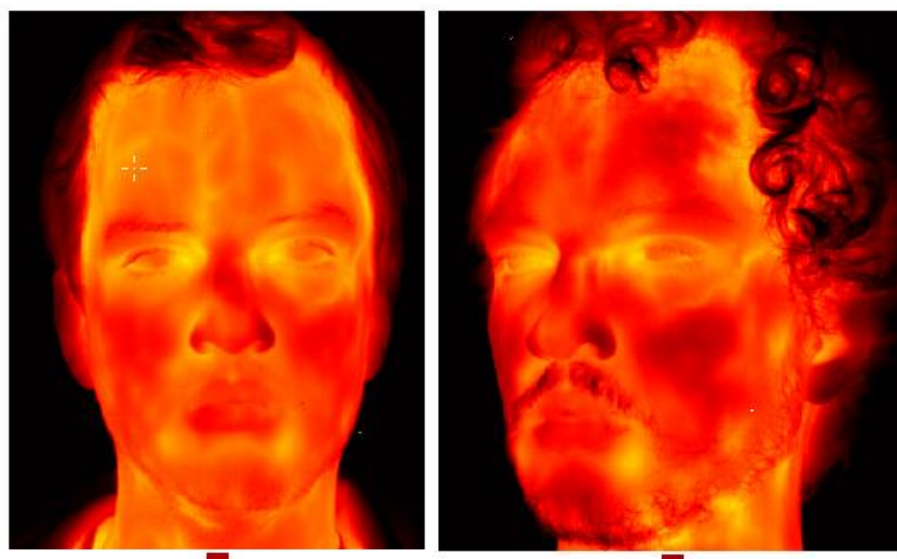


Figure 5: A sample of UL-FMTV after time-lapse: (a) Room temperature, and (b) after physical activity

**4.2 How to obtain the database and use it**

The database is available for public use for research purposes and non-commercial use only. The database can be accessed and downloaded from the website of QIRT. There is also a Matlab file attached to the database which helps the researcher to easily read the database files in both single frame or video format. The reader may contact the corresponding author for more details.

## 5. Conclusions

In this paper, we presented Université Laval Face Motion and Time-Lapse Video Database (UL-FMTV), the largest facial video database of the world in the Mid-Wave Infrared spectrum. The lack of a large video database with high resolution gathered during multiple sessions within a relatively large period which contains images of subjects with different ethnicities, ages and sexes are missing in the literature. In this database, we not only focused on these factors but also considered a wide range of facial poses and expressions. The database is obtained from 238 subjects and is available for public use for research purposes only. This database consists of high-resolution thermal videos of subjects within a time period of four years, which makes it very interesting to extract superficial under-skin tissues such as blood vessels. This videos also make it possible to train statistical face models such as AAMs which has applications such as face tracking, facial expression normalization, and facial pose estimation.
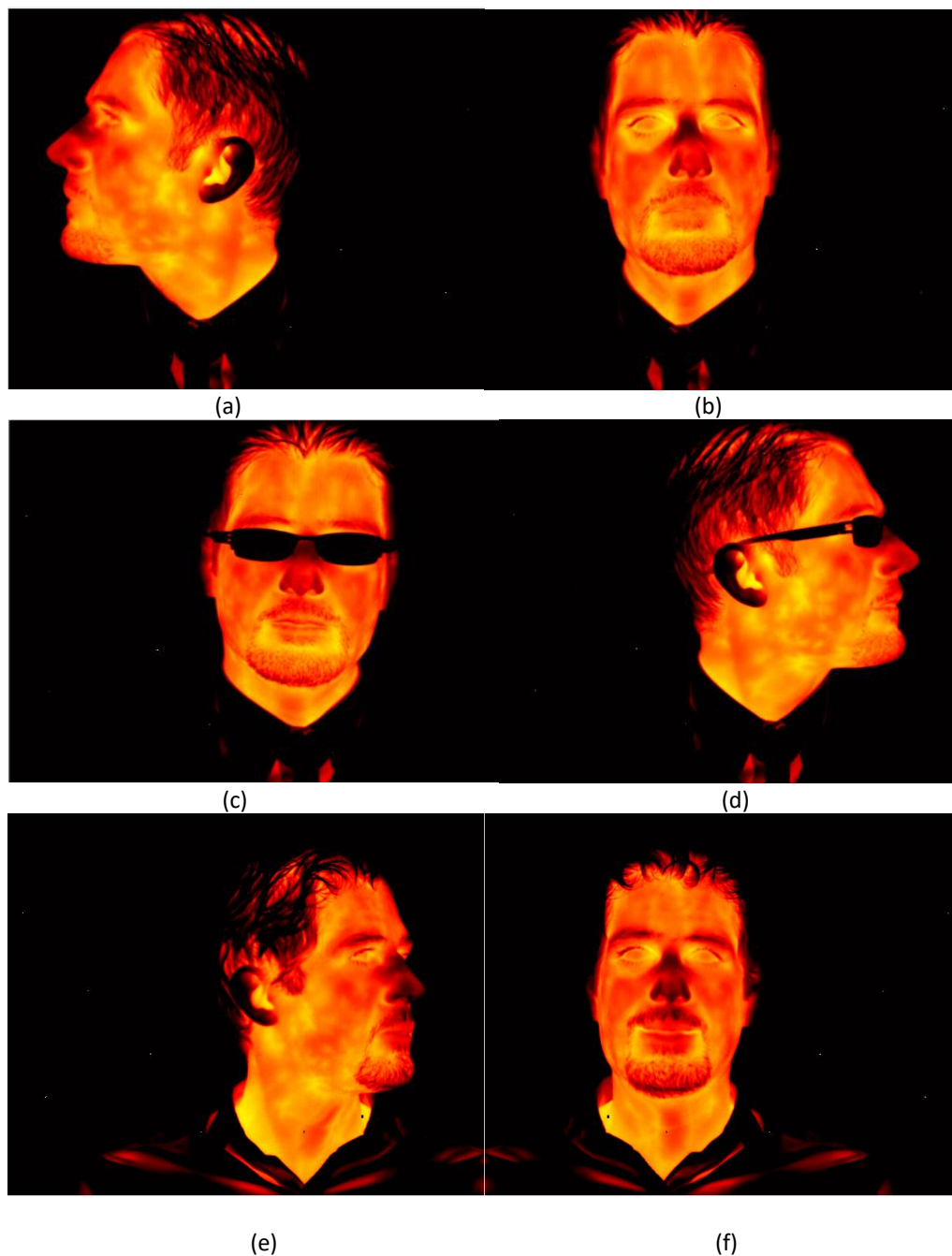
## 6. Acknowledgment

Figure 6:   A sample of UL-FMTV after time-lapse: (a) extreme pose, (b) frontal, (c) glasses, (d) glasses + extreme pose, (e) after 23 months with pose and (f) after 23 months frontal view
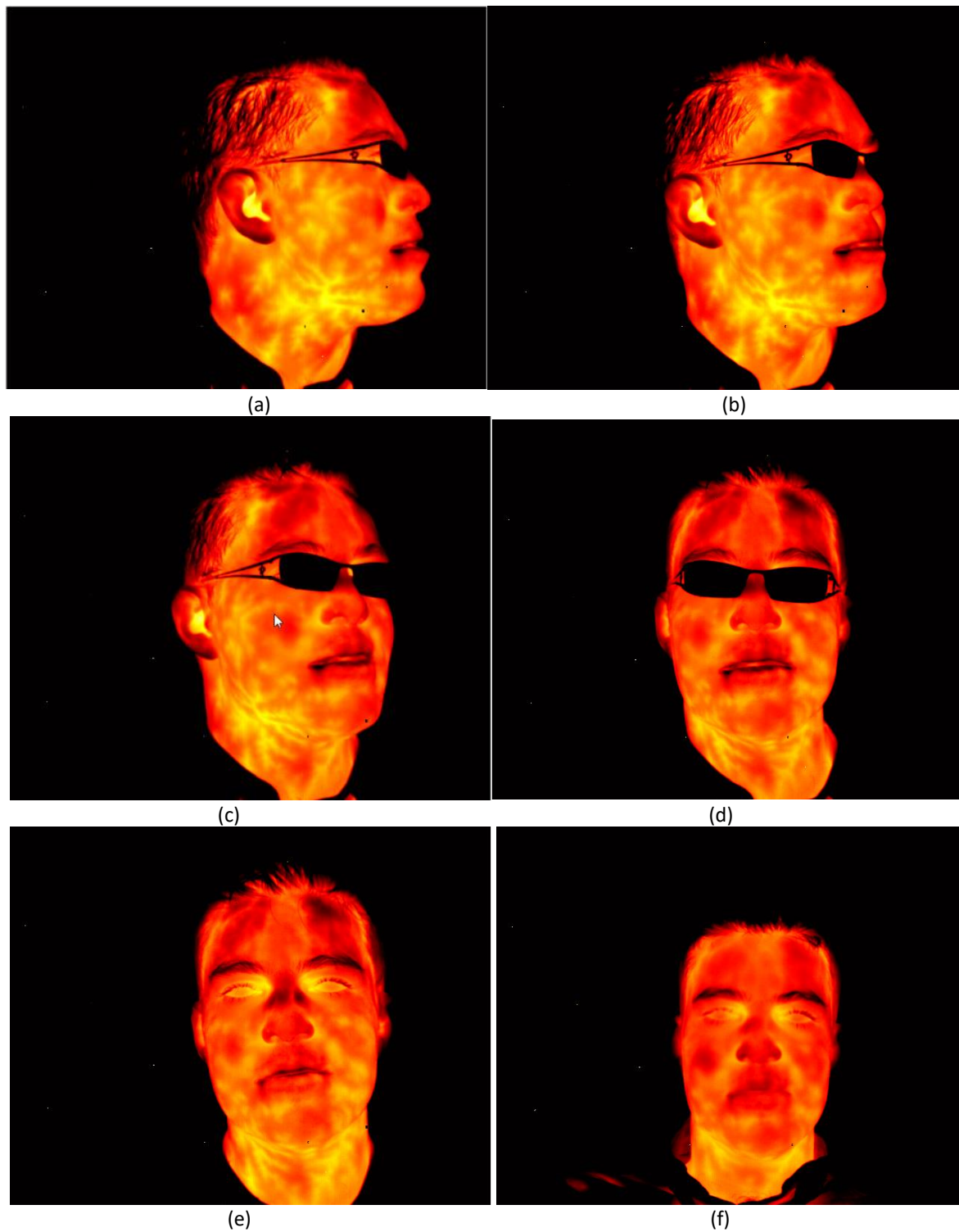
Figure 7: Another sample of UL-FMTV after time-lapse: (a) extreme pose+glasses, (b) pose+glasses, (c) mid-profile+glasses, (d) frontal+glasses, (e) frontal, and (f) after 23 months frontal view

(1)

**REFERENCES**

[1]   R. Gross, I. Matthews, and S. Baker. Active appearance models with occlusion. Image and Vision Computing (special issue on Face Processing in Video), 1(6) :593–604, 2006.

[2]   V. Blanz and T. Vetter. Face recognition based on fitting a 3D morphable model. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 25(9) :1063–1074, 2003.

[3]   U. Mohammed, S. Prince, and J.. Kautz. Visio-lization : generating novel facial images. ACM Transactions on Graphics (TOG), 28(3) :57 :1–57 :8, 2009.

[4]   M. Nishiyama and O. Yamaguchi. Face recognition using the classified appearance-based quotient image. In Proc. IEEE International Conference on Automatic Face and Gesture Recognition (FG), pages 49–54, 2006.

[5]   L. Wolf and A. Shashua. Learning over sets using kernel principal angles. Journal of Machin Learning Research (JMLR), 4(10) :913–931, 2003.

[6]   S.Z. Li et al., "AuthenMetric F1: A Highly Accurate and Fast Face Recognition System," Proc. Int'l Conf. Computer Vision, Oct. 2005.

[7]   R. S. Ghiass, A. Bendada, and X. Maldague. Infrared face recognition : A review of the state of the art. In Proc. International Conference on Quantitative Infrared Thermography (QIRT), pages 533–540, 2010.

[8]   P. Buddharaju, I. Pavlidis, and P. Tsiamyrtzis, "Phsiology-based face recognition using the vascular network extracted from thermal facial images: A novel approach ", Proceedings of the IEEE International Conference on Advanced Video and Signal based Surveillance , Como, Italy, September 15-16, 2005.

[9]   P. Buddharaju, I.T. Pavlidis, P. Tsiamyrtzis and M. Bazakos, "Physiology-Based face Recognition in the Thermal Infrared Spectrum", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29, No. 4, pp. 613-626, April 2007.

[10]  WU, S., SONG, W., JIANG, L.J., XIE, S., PAN, F., YAU, W.Y., AND RANGANATH, S., "Infrared face recognition by using blood perfusion data," In Proceedings of the Intl Conf Audio Video Based Biometric Person Authentication (New York, USA., July 20-22, 2005.) Pp. 320-328.

[11]  Ghiass, Reza Shoja, et al. "A unified framework for thermal face recognition." *International Conference on Neural Information Processing.* Springer, Cham, 2014.

[12]  Ghiass, Reza Shoja, et al. "Illumination-invariant face recognition from a single image across extreme pose using a dual dimension AAM ensemble in the thermal infrared spectrum." *Neural Networks (IJCNN), The 2013 International Joint Conference on.* IEEE, 2013.

[13]  T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In Proc. European Conference on Computer Vision (ECCV), 2 :484–498, 1998.

[14]  I. Matthews and S. Baker. Active appearance models revisited. International Journal of Computer Vision (IJCV), 60(2) :135–164, 2004.

[15]  S. Romdhani and T. Vetter. Efficient, robust and accurate fitting of a 3D morphable model. In Proc. IEEE International Conference on Computer Vision, pages 59–66, 2003.

[16]  Ghiass, Reza Shoja, et al. "Infrared face recognition: A comprehensive review of methodologies and databases." *Pattern Recognition* 47.9 (2014): 2807-2824.

[17]  Ghiass, Reza Shoja. *Face Recognition Using Infrared Vision.* Diss. Université Laval, 2014.