A multispectral approach to robust human skin detection

Moritz Störring, Hans J. Andersen, and Erik Granum
Computer Vision and Media Technology Laboratory, Aalborg University
Aalborg, Denmark
Email: mst,hja,eg@cvmt.dk

Abstract

Using multiple wave-bands is common in multispectral imaging and remote sensing, e.g., to enable improved reproduction of colours or to detect the crop status in a field, respectively. Computer vision methods, however, mainly rely on monochrome or RGB wave-bands, although segmentation tasks might be improved considerably using more or other wave-bands than the standard RGB.

This paper presents and investigates a new approach for detecting human skin using a combination of standard RGB bands and three near infrared bands. Simulations under changing illumination conditions and a preliminary experiment show an improved robustness over pure RGB based approaches.

Introduction

Robust face and hand tracking is important for applications such as face and gesture recognition in interfaces for HCI (human computer interaction). An often used approach to do unobtrusive tracking is computer vision. Robust and reliable segmentation of human skin is then crucial for the success of such a system. The segmentation may be done using cues like colour, motion, or shape. Skin colour detection is more and more used [1] because it is invariant to size, shape, and viewing direction. However, other materials with different spectral compositions than skin may have the same RGB colour components as skin, i.e. metamerism [2], and hence result in false detections. Furthermore, changes in the illumination spectrum may cause changes in the light reflected by skin, i.e. changes in the apparent skin colour, which may also cause false positives or negatives [3].

These problems might be overcome using other wave-bands than the three RGB bands or more bands, which is commonly done in other fields like multispectral imaging or remote sensing. Multispectral imaging got much attention in the past decade for accurate reproduction, archiving or food-inspection using 30 bands or even more in the visible spectrum. This is usually done under controlled illumination and environment conditions. A series of workshops has been held on multispectral imaging [4] and various end user products are becoming available, e.g. Sony recently released a four colour CCD [5] with two bands in the green wavelengths for better colour reproduction.

Despite the success of using multiple wave-bands in multispectral imaging and remote sensing, this has not yet found much attention in the computer vision community except for some few special applications, e.g. agricultural applications [6, 7]. Approaches using more than one wave-band are usually based on the RGB bands of colour cameras, often consumer quality cameras, e.g. webcams.

Angelopoulou et. al. [2] investigated the spectral reflectance of skin and presented a multispectral skin colour modelling approach using five Gaussian distributions. They suggested to capture skin reflectance with five bandpass filters assuming the illumination is known. In [8] it is shown that skin reflectance spectra may be used not only to detect, but also to recognise individual people. Other approaches for detecting skin are using near infrared (1-3μm) and far infrared (5-8μm) cameras, respectively [9, 10]. However, these cameras have low resolutions and are rather expensive because they cannot use standard silicon technology which is sensitive only up to wavelengths of 1.1μm (see dashed curve in figure 3).

In this paper a multispectral approach for skin detection under multiple illuminations is presented. In particular, a combination of standard RGB bands with three NIR (near infrared) bands below 1μm is investigated.

The next section briefly reviews the theory used in this paper. Then the proposed method and simulation results are presented, which is followed by an initial experiment. Finally the results and the applicability are discussed.

Background

The reflected light from a material may be described with the Dichromatic Reflection Model [11] which states that the reflected light is the sum of the material's body and surface reflections. Surface reflections are taking place directly at the material's surface without penetrating into it. They have approximately the same spectral composition as the light source and under certain illumination and viewing geometries they give highlights. These are not considered in the following since they usually only occupy small areas and they may be detected separately [11, 12]. Body reflections are due to light penetrating into the mate-

rial where it is wavelength dependently absorbed and scattered around. This reflection gives persistent spectral information about the material, e.g. its colour.

The image formation process may be modelled by spectral integration as in eq. 1. C_i are the outputs of the camera channels, i.e. for a colour camera $i \in \{R, G, B\}$.

$$C_i = G_i \int E(\lambda)\rho(\lambda)f_i(\lambda)d\lambda \quad (1)$$

where λ is the wavelength, E the illumination spectrum, ρ the material's reflectance, and f_i the spectral sensitivities (examples for E, ρ, f_i are illustrated in figures 1, 2, and 3, respectively). G_i depends on the camera, and the illumination- and viewing-geometry (photometric angles). The output of a camera for a given reflectance ρ, thus depends on G_i and the illumination's spectral composition and amplitude (intensity). In order to get invariant to G_i and the intensity one may transform the RGB values to another colour space [1, 13, 3], e.g. normalised RGB or band ratios. Band ratios are used in the following and may be defined as $r = C_R/C_B$ and $g = C_G/C_B$. Band ratios are not invariant to changes in the illumination's spectral composition.

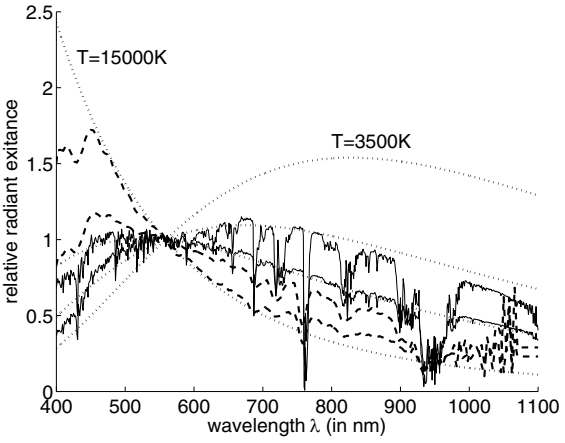

Figure 1: *Spectra of different illuminants normalised at $\lambda = 560nm$. Blackbody 3500-15000K (dotted), daylight simulated with SMARTS2 [14] (solid), and daylight from the Joensuu database [15] (dashed).*

Marchant and Onyango [19, 20] showed that for certain illuminant families a function F of the band ratios of a material's reflections exists that is invariant to illumination changes within one illuminant family:

$$F = \frac{r}{g^A} \quad (2)$$

where the exponent A depends on the centre wavelengths of the sensitivities f_i. Hence, A is constant for a given camera and may be precalculated. F is invariant to illumination changes within illuminant families that may be described by

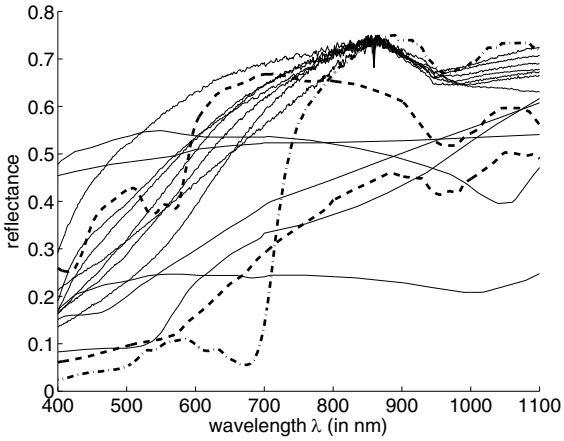

Figure 2: *Reflectance spectra of skin and other objects in visible and NIR wavelengths. Dashed thick lines for light skin (upper) and dark skin (lower) [16], dash-dotted for green vegetation [17], and thin solid lines for, e.g., different woods [15], reddish brick, and reddish metals [18].*

$$E(\lambda, S) = a(\lambda) + u(\lambda)b(S) \quad (3)$$

where $a(\lambda)$ and $u(\lambda)$ are any function of λ but not S and $b(S)$ is any function of S, but not λ [20]. One family of such illuminants are Blackbody radiators, which have the following form, when represented by the Wien approximation [18]: $E(\lambda, T) = c_1\lambda^{-5}\exp(-c_2/T\lambda)$, where T is the colour temperature and the parameter S in $b(S)$ which F is invariant to. In [20] it was shown that the family of daylights may also be described by eq. 3. It should be noted again that F is different for different illuminant families.

The exponent A may be calculated from the centre wavelengths of the sensitivities. Another possibility, which is chosen here, is to simulate the camera outputs using eq. 1 for different illuminants. Taking the logarithm eq. 2 gives

$$\log(r) = A\log(g) + \log(F) \quad (4)$$

Hence, A may be estimated by fitting a line into the simulated data.

Skin Detection

In this section F is simulated for skin and examples of other materials under varying illumination conditions. This is first done in the visual spectrum using the spectral sensitivities f_i of a RGB camera to simulate F_{VIS}. Then three sensitivities in the NIR spectrum are introduced and F_{NIR} is simulated. Finally, F_{VIS} and F_{NIR} are combined.

Light sources are often characterised by their spectral composition. General purpose light sources, such as daylight, electric light bulbs, and fluorescent lamps, have a

non-zero spectrum in the visible wavelengths. Some light sources like daylight and electric light bulbs also have a non-zero spectrum in the NIR wavelengths. Figure 1 shows Blackbody (e.g. electric light bulb) and daylight spectra in the visible and NIR wavelengths. However, some artificial light sources block the NIR radiation, e.g. fluorescent lamps. The latter are therefore not be considered further.

In [21] it was shown that most general purpose light sources may be approximated with Blackbody radiators when seen by a RGB camera. In the following the illumination will be modelled by Blackbody radiators with colour temperatures from 3500-15000K. It should, however, be noted here that this approximation might not be appropriate in NIR.

Figure 2 shows reflectance spectra of skin and examples of other materials. Non-skin reflectance spectra of 17 materials were chosen that have reflectance properties close to skin (metamer) when viewed by a RGB camera and/or when viewed in the NIR wavelengths, respectively. In the visible spectrum, e.g. aspen, birch, polished rose granite are rather close to the reflectance of skin and in the NIR, e.g. vegetation is very similar to skin reflectance.

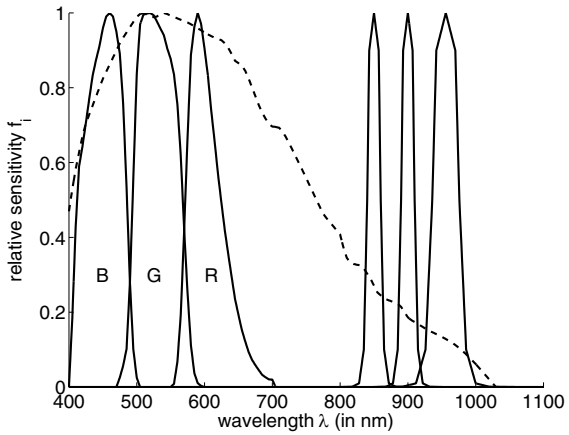

Figure 3: *Spectral sensitivities. Solid lines: RGB sensitivities (400-700nm) of JAI CV-M90 and three sensitivities in the NIR wavelengths (800-1000nm). Centre wavelengths are 850, 900, and 950nm. Dashed line: JAI CV-M4+.*

Visible Band

The RGB outputs of a camera are simulated by spectral integration (eq. 1) using the RGB sensitivities in figure 3, the reflectance spectra in figure 2, and varying illumination conditions with colour temperatures from T=3500-15000K and ΔT=100K. F_{VIS} is calculated using the band ratios $r = C_R/C_G$ and $g = C_B/C_G$. Figure 4 shows the simulation result of F_{VIS}.

It can be seen that several materials fall together with skin, e.g. aspen and birch. This would result in false positives and/or negatives when using a colour based skin/non-skin classifier. It should be noted that this is not a particular problem for band ratios but also the case in other colour spaces, see e.g. [13].

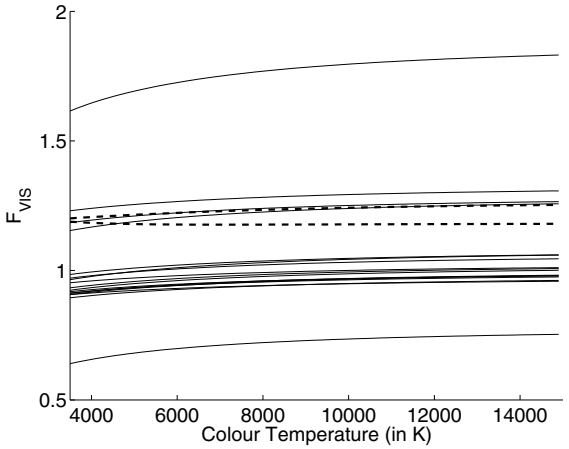

Figure 4: *F_{VIS} as function of the colour temperature. Dashed thick lines for skin, thin solid lines for other materials.*

NIR Band

The spectral information in the NIR band may be exploited to distinguish skin from materials with similar F_{VIS} values, i.e. use of additional sensitivities in NIR. The reflectance curves of skin have a characteristic local minimum at around λ = 950nm, figure 2, that is due to the high water content in skin. Also other biological material have local minima around that wavelength, however only vegetation has a similarly pronounced minimum.

A function F_{NIR} might be introduced using the same approach as above for the visible band (F_{VIS}). This requires three bands in the NIR in order to get two ratios $r_{NIR} = C_{NIR1}/C_{NIR2}$ and $g_{NIR} = C_{NIR3}/C_{NIR2}$. These bands might be found iteratively by maximising ΔF_{NIR} between skin and other materials. Here it was chosen to fix one sensitivity to the local minimum at λ = 950nm and only iterate over two sensitivities. Due to the sensitivity range of silicon technology (300-1100nm, dashed line in figure 3) they had to be below 950nm. Furthermore, the sensitivities were modelled as standard optical bandpass filters, e.g. available from Andover corp., NH, USA.

The sensitivities are shown in figure 3. The simulation result is shown in figure 5. It can be seen that the F_{NIR} of skin is rather close to that of vegetation.

Combining Visible and NIR

Figure 6 shows F_{VIS} as function of F_{NIR}. It can be seen that the skin distributions are now separated from the other materials' distributions. In this simulation additive normal zero mean distributed noise with a signal to noise ratio

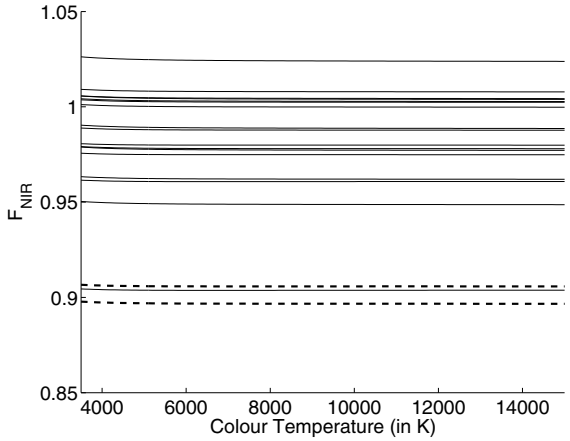

Figure 5: F_{NIR} *as function of the colour temperature. Dashed thick lines for skin, thin solid lines for other materials.*

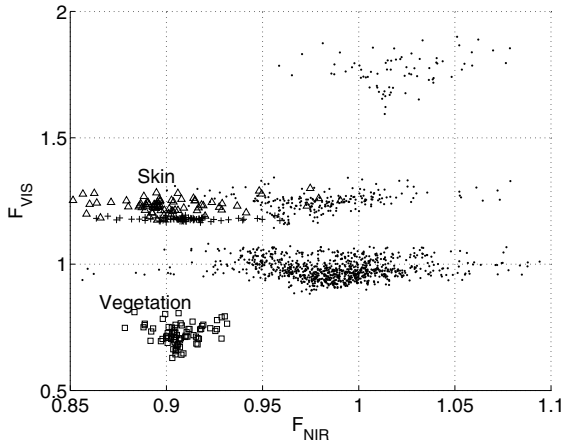

Figure 7: F_{VIS} *as function of* F_{NIR} *simulated with 60dB SNR. Light skin (+), dark skin (△), other materials (·).*

(SNR) of 68dB was added. The SNR definition in eq. 5 was used.

$$SNR = 20\log_{10}\left(\frac{S}{N}\right) \tag{5}$$

The JAI CV-M4+ (figure 3) has according to the datasheet a SNR >57dB. Figure 7 shows the same simulation but with SNR = 60dB. Now the aspen and skin distributions start to overlap, which is due to F_{NIR}.

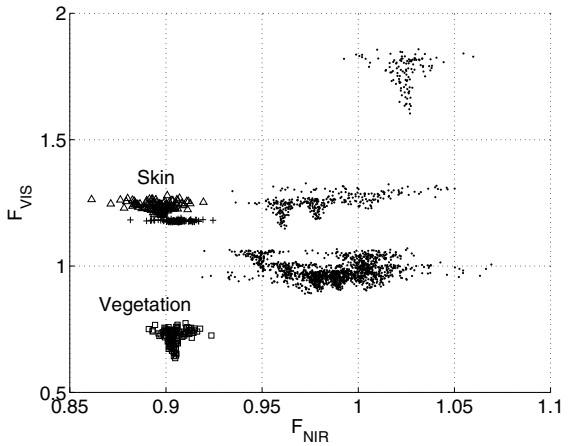

Figure 6: F_{VIS} *as function of* F_{NIR} *simulated with 68dB SNR. Light skin (+), dark skin (△), other materials (·).*

NIR Experiment

This section shows a preliminary experiment in the NIR band in order to verify whether the local minimum of skin reflectance at λ=950nm is big enough to be detected with a state of the art camera. Images of skin and every day objects such as wood, fruit, water, cloth, plastic, vegetation etc. were captured using two optical bandpass filters with centre wavelengths λ_{C1}=830nm and λ_{C2}=950nm, respectively, and bandwidths λ_{BW1}=20nm and λ_{BW2}=40nm, respectively. Figure 8 shows an example taken with a 830nm filter. In figure 9 the ratio between the λ_{C1}=830nm and λ_{C2}=950nm images was taken and thresholded. As expected from the simulation, materials with high water content may not be removed in the NIR, however, wooden and other materials are removed.

Figure 8: *Image with human skin and other every day materials captured under tungsten light with a JAI CV-M4+ camera and an optical bandpass filter with* λ*=830nm centre wavelength and 20nm bandwidth.*

Discussion

This paper suggested and investigated the use of standard RGB bands combined with three NIR bands for skin detection under changing illumination conditions. Simulations with reflectance spectra of skin and other materials with reflectance characteristics similar to skin show promising results.

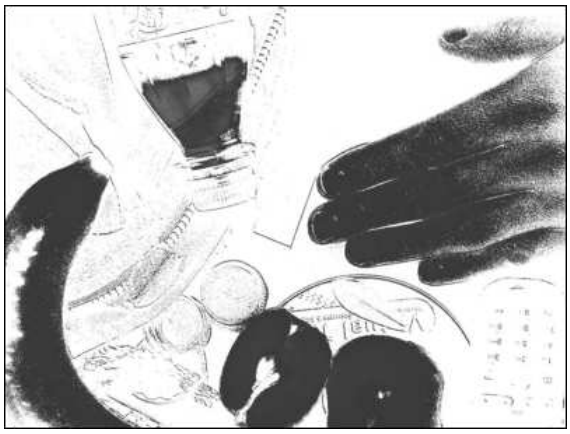

Figure 9: *Thresholded ratio between two images captured with filters of λ=830nm and λ=950nm, respectively, centre wavelength.*

Furthermore, preliminary tests using only the ratio between two real images of skin and other materials captured with bandpass filters of 830nm and 950nm, respectively, show that the local minimum around 950nm might be used as a feature for skin detection when using blackbody illumination.

The approximation of light sources with blackbody radiators has been proven to work in the visible band and it is also a good approximation in NIR for tungsten light bulbs. However, daylight spectra deviate considerably from blackbody radiators in the NIR, particularly around the local minimum of skin. It has to be verified whether the proposed NIR sensitivities also work for other light sources such as daylight, figure 1. This might be done using SMARTS (Simple Model of Atmospheric Radiative Transfer of Sunshine) [14] to simulate daylight.

Currently 17 materials have been tested. A more extensive simulation with spectra of more materials is necessary. In the visible spectrum large databases exist, e.g. [22], which could be used. In the NIR there is, to the knowledge of the authors, less data available.

As already shown by the four colour Sony CCD, new generations of imaging sensors could have more than the standard RGB sensitivities, e.g. in the NIR band and by that enabling robust skin detection.

Acknowledgements

This research is in part funded by the ARTHUR project (IST-2000-28559) under the European Commissions IST program. This support is gratefully acknowledged.

References

1. Erik Hjelmas and Boon Kee Low, Face detection: A survey, Computer Vision and Image Understanding, 83(3), 236 (2001).
2. Elli Angelopoulou, Rana Molana and Kostas Daniilidis, Multispectral skin color modeling, in IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, Hawaii, vol. 2, pp. 635–642 (2001).
3. Moritz Störring, Hans Jørgen Andersen and Erik Granum, Physics-based modelling of human skin under mixed illuminants, Journal of Robotics and Autonomous Systems, 35(3–4), 131 (2001).
4. Homepage for multispectral imaging, http://www.multispectral.org.
5. Sony-Corporation, Realization of natural color reproduction in digital still cameras, closer to the natural sight perception of the human eye; technology development to realize 4-color filter ccd and new image processor, achieving halving of color reproduction error, Press Release (2003).
6. Matthias F. Carlsohn, Guest editorial: Spectral imaging for real-time imaging applications, Real-Time Imaging (Special Issue on Spectral Imaging), 9(4), 229 (2003).
7. Hans Jørgen Andersen, Christine M. Onyango and John A. Marchant, The design and operation of an imaging sensor for detecting vegetation, Journal of Imaging Science and Technology, 11, 144 (2000).
8. Zhihong Pan, Glenn Healey, Manish Prasad and Bruce Tromberg, Face recognition in hyperspectral images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(12), 1552 (2003).
9. Jonathan Dowdall, Ioannis Pavlidis and George Bebis, Face detection in the near-ir spectrum, Image and Vision Computing, 21(7), 565 (2003).
10. Christopher K. Eveland, Diego A. Socolinsky and Lawrence B. Wolff, Tracking human faces in infrared video, Image and Vision Computing, 21(7), 579 (2003).
11. Gudrun J. Klinker, Steven A. Shafer and Takeo Kanade, A physical approach to color image understanding, Int. Journal of Computer Vision, 4(1), 7 (1990).
12. Hans Jørgen Andersen and Moritz Störring, Classifying body and surface reflections using expectation-maximisation, in PICS: Image Processing, Image Quality, Image Capture Systems Conference, Rochester, New York, USA, pp. 441–446 (2003).
13. J. Birgitta Martinkauppi, Maricor N. Soriano and Mika H. Laaksonen, Behavior of skin color under varying illumination seen by different cameras at different color spaces, in M. A. Hunt, editor, SPIE Machine Vision in Industrial Inspection IX, San Jose, California, USA, vol. 4301 (2001).
14. Christian A. Gueymard, Parameterized transmittance model for direct beam and circumsolar spectral irradiance, Solar Energy, 71(5), 325 (2001).
15. Jussi Parkkinen and Pertti Silfsten, Spectra databases, http://www.cs.joensuu.fi/~spectral/index.html.
16. R. R. Anderson, J. Hu and J. A. Parrish, Optical radiation transfer in the human skin and applications in *in vivo* remittance spectroscopy, in R. Marks and P. A. Payne, editors, Bioengineering and the Skin, MTP Press Limited, chap. 28, pp. 253–265 (1981).
17. Terence P. Dawson, Paul J. Curran and Stephen E. Plummer, LIBERTY - modeling the effects of leaf biochemical concentration on reflectance spectra, Remote Sensing of Environment, 65(1), 50 (1998).
18. Günter Wyszecki and W. S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, John Wiley and Sons (1982).
19. J.A. Marchant and C.M. Onyango, Shadow-invariant classification for scenes illuminated by daylight, Journal of the Optical Society of America A, 17(11), 1952 (2000).
20. J. A. Marchant and C. M. Onyango, Color invariant for daylight changes: relaxing the constraints on illuminants, Journal of the Optical Society of America A, 18(11), 2704 (2001).
21. Graham D. Finlayson and Gerald Schaefer, Solving for colour constancy using a constrained dichromatic reflection model, Int. Journal of Computer Vision, 42(3), 127 (2001).
22. Graphic technology – Standard object colour spectra database for colour reproduction evaluation (SOCS), ISO/TR 16066:2003 (2003).

Biography

Moritz Störring is assistant professor at the Computer Vision and Media Technology Laboratory (CVMT), Aalborg University, Denmark. He studied Electrical Engineering at the Technical University of Berlin, Germany and at the Institut National Polytechnique de Grenoble, France. He graduated in 1998 in Berlin. From 1998 to 2001 he was a research assistant at CVMT within the European TMR project SMART II. His research interests include colour vision, outdoor computer vision, face and skin detection, vision based HCI, and augmented reality.

Hans J. Andersen has received a Ph.D. degree from the Computer Vision and Media Technology Laboratory, Aalborg University within the field Outdoor Computer Vision. After his Ph.D. he has worked as Research and Development Manager at Eco-Dan A/S introducing the first commercial system for computer vision supported guidance of agricultural machinery and as a researcher with the Technology and Innovation Department at Siemens Mobile Phones. He is now assistant professor at the CVMT.

Erik Granum is professor of information systems and head of CVMT, Computer Vision and Media Technology Laboratory, and co-founder of VR Media Lab at Aalborg University, Denmark. He is and has been co-ordinator of several national and international research projects and networks in computer vision and virtual reality and partner of many others. His research interests cover pattern recognition, continually operating vision systems, colour vision, vision guided multimedia interfaces, visualisation, virtual reality, and autonomous agents.