Chapter 6
Color and Thermal Image Fusion for Augmented Reality in Rescue Robotics

Ludek Zalud, Petra Kocmanova, Frantisek Burian and Tomas Jilek

Abstract At the beginning of this article, the authors address the main problems of todays remotely-operated reconnaissance robots. The reconnaissance robots Orpheus-AC, Orpheus-AC2 and Orpheus-Explorer, made in the Department of Control and Instrumentation (DCI), are then shortly described. Since all the described robotic systems use visual telepresence as the main control technique, visual information from the robots surroundings is essential for the operator. For this reason, the authors make a fusion of data from a Charge-Coupled Device (CCD) color camera, and a thermovision camera to provide the operator with data in all visibility conditions, such as complete darkness, fog, smoke, etc.

Keywords Robot · User interface · Telepresence · Augmented reality

6.1 Introduction

The reconnaissance of dangerous areas is one of the most challenging tasks for todays robotics. According to many indications, e.g. from the Robocup Rescue League community where the DCI team is involved [1, 2], it seems that nowadays

L. Zalud (✉) · P. Kocmanova · F. Burian · T. Jilek
CEITEC Central European Institute of Technology, Technicka 10, 61600 Brno, Czech Republic
e-mail: ludek.zalud@ceitec.vutbr.cz

P. Kocmanova
e-mail: petra.kocmanova@ceitec.vutbr.cz

F. Burian
e-mail: frantisek.burian@ceitec.vutbr.cz

T. Jilek
e-mail: tomas.jilek@ceitec.vutbr.cz

H. A. Mat Sakim and M. T. Mustaffa (eds.), *The 8th International Conference on Robotic, Vision, Signal Processing & Power Applications*,
Lecture Notes in Electrical Engineering 291, DOI: 10.1007/978-981-4585-42-2_6,

the development of practical and usable reconnaissance robots [3] is aimed at the following tasks:

- A larger number of robots controlled by one operator.
- Easy and intuitive human-to-robot interface.
- For many kinds of reconnaissance missions it would be highly beneficial if the user interface would somehow emphasize alive people.

The authors propose a possible solution of the abovementioned problems through an advanced user interface program called CASSANDRA and show its application on Orpheus reconnaissance robots. Three different robots: Orpheus-AC, Orpheus-AC2 and Orpheus-Explorer, completely built in our laboratory, are roughly described in the Sect. 6.2, providing an example of remotely controlled robotic systems possibly controlled by CASSANDRA, while the described data-fusion is nowadays present on Orpheus-Explorer.

6.2 Orpheus Robots

Orpheus robots have been developed at Department od Control and Instrumentation (DCI) since 2003. The first version was called simply Orpheus, and our team was quite successful in Robocup Rescue 2003 world competition in Padova, Italy—we won the competition (see [2]). In 2003–2006 we improved/rebuilt the robot to the version Orpheus-X2 (see [4]). In 2006 we were asked to make a military version of the robot. The prototype was finished in 2007 and named Orpheus-AC (Army and Chemical) [5]. In 2009 we started development of second generation, based on Orpheus-A2 platform. We decided to make two basic modifications—Orpheus-AC2 for chemical and nuclear contamination measurements and Orpheus-Explorer for more general reconnaissance missions and victim search (Fig. 6.1).

The hereinafter described fusion is done on Orpheus-Explorer robot, which is equipped with novel sensory head containing wide-FOV (Field of View) overview camera AXIS 1114 with 2.3 mm lens, wide FOV thermal camera NEC C100C with 8 mm lens, and high-resolution reconnaissance camera AXIS Q1755 with "night vision" IR mode, see Fig. 6.2.

6.3 Visible Spectrum and Thermovision Data Fusion

The goal of the research described in this subsection is to improve the user interface to be a system that: makes robot control possible virtually in all visibility conditions [6, 7, 8, 9], such as fog, smoke or darkness; displays the most appropriate data or fused data in a convenient and intuitive way, visually emphasizes living victims, permits the use of the same data for digital map building and self-localization. This is done through data-fusion from CCD color camera and thermal imager [10, 11].

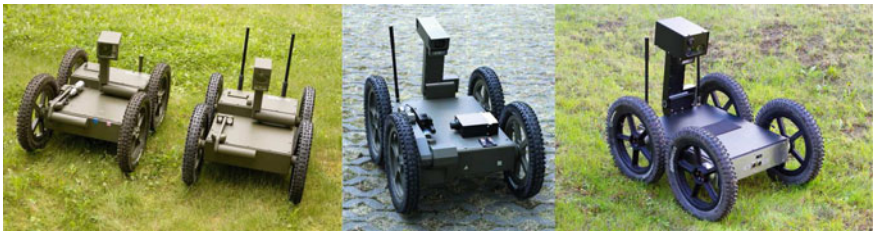

Fig. 6.1 Orpheus robots (*from left*): Orpheus-AC prototype, Orpheus-AC, Orpheus-AC2, Orpheus-Explorer

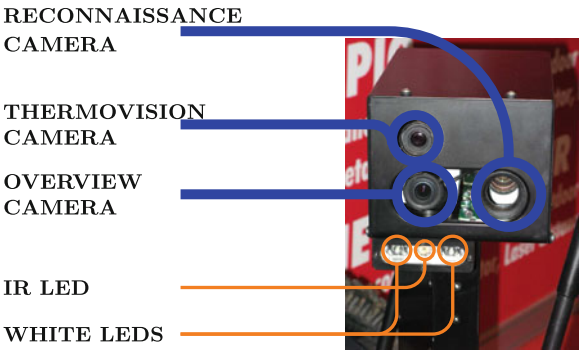

Fig. 6.2 Orpheus-Explorer sensory head

The basic problem of this data fusion is that the goal is to mix the data from the visible spectrum sensed by a color CCD camera and LWIR (long wave infrared) spectrum measured by the thermovision imager to the visible spectrum shown on a display [12, 13]. Since the spectrum to be displayed is evidently wider than the one we have at our disposal, it seems this task may not be done without some kind of compression of the color data.

We will use HSV color model to describe the colors. If we express a CCD camera image in the HSV mode [14] and make histograms of it, we clearly see that the distribution of S and V is not balanced and there are few (if any) pixels with $S = 1$ and $V = 1$. This is a result of the restrictions of CCD and their data processing.

From this we can conclude that if we add pixels with colors that have an arbitrary H parameter and $S = 1$ and $V = 1$, they will thus clearly be perceptible by the operator. In other words, it could be said we are adding a rainbow of colors to the ordinary image, which while perhaps uncommon (or even present) in an image, may well be used to emphasize parts of the image.

Regarding the previous text, the whole procedure of data mixing described as follows:

- The image from a thermovision camera is digitalized, if it is not already digital.
- The temperatures that are not near human body temperatures are filtered off.

- The pixels that are to be displayed are recalculated—H corresponds to frequency, S = 1 and V = 1.
- This thermovision image is rendered (e.g. alpha-blended) over the CCD camera image.

So we can write the color transformation:

$$H' = \frac{H}{60} = 6 \times (1 - R_{TH}) \tag{6.1}$$

$$H_A = H' - \lfloor H' \rfloor \tag{6.2a}$$

$$H_B = 1 - H_A \tag{6.2b}$$

$$(R, G, B) = \begin{cases} (1, H_A, 0) & \text{if } \lfloor H' \rfloor = 0 \\ (H_B, 1, 0) & \text{if } \lfloor H' \rfloor = 1 \\ (0, 1, H_A) & \text{if } \lfloor H' \rfloor = 2 \\ (0, H_B, 1) & \text{if } \lfloor H' \rfloor = 3 \\ (H_A, 0, 1) & \text{if } \lfloor H' \rfloor = 4 \\ (1, 0, H_B) & \text{if } \lfloor H' \rfloor = 5 \\ (0, 0, 0) & \text{otherwise} \end{cases} \tag{6.3}$$

Now, we have to remove the pixels that do not correspond to the temperatures, we are interested in 6.4.

$$(R, G, B) \cdot (0.30, 0.59, 0.11) < K, \tag{6.4}$$

where K ia a constant that corresponds to a temperature. In our case we do not have so called "calibrated" thermal camera, but we have made our own calibration using Peltier, and K corresponds approximately to human body temperature, while the parameter is user selectable, even during mission.

The CCD camera and thermovision camera images spatial correspondence scaling is yet to be mentioned and still to be done. This results from the fact that the field-of-view of the cameras is different, they have different positions, they are not aligned to be perfectly parallel, and each of the sensors has a different resolution.

6.3.1 CCD and Thermocamera Alignment

In the Fig. 6.3, the problem of CCD camera and thermovision camera misalignment is defined, with all the important features.

All the important parameters of the two-camera-system are summarised in the Table 6.1. For these calculations and experiments we use AXIS M1114 CCD network camera with Computar 2.3 mm DC lens, and NEC C100C thermal camera OEM module with 8 mm lens inserted in Orpheus-Explorer.

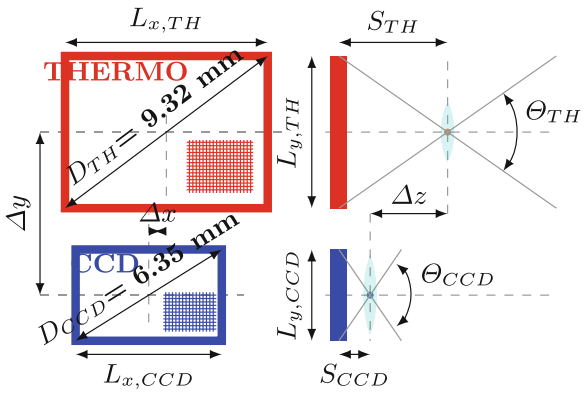

Fig. 6.3 Cameras position, resolution and chip size—parameter definition

Table 6.1 Parameters of both sensor chips

	Resolution				Chip size		FOV	
	D [mm]	S [mm]	R_x [px]	R_y [px]	L_x [mm]	L_y [mm]	Φ [°]	Θ [°]
Thermo	9.32	8.0	320	240	7.46	5.60	50.0	37.5
CCD	6.35	2.3	1280	800	5.38	3.36	98.9	72.3

In the Table 6.1 the chip diameter D, focal length S, and chip resolution R are known from specification of cameras. The chip sizes L_x and L_y were calculated from camera specifications by simple formula 6.5.

$$L_x = \frac{R_x \cdot D}{\sqrt{R_x^2 + R_y^2}}, \quad L_y = \frac{R_y \cdot D}{\sqrt{R_x^2 + R_y^2}} \tag{6.5}$$

Next columns in the table, the field of view (Θ for the vertical axis, and Φ for the horizontal axis) of the cameras can be calculated using Eq. 6.6. The L_x and L_y parameters in these equations contain appropriate dimensions of image sensor chip, solved previously.

$$\tan\left(\frac{\Phi}{2}\right) = \frac{L_x}{2 \cdot S}, \quad \tan\left(\frac{\Theta}{2}\right) = \frac{L_y}{2 \cdot S} \tag{6.6}$$

Suppose that the camera system with optics is linear. In this case, we can draw Fig. 6.4. The observed point P placed at position (P_x, P_y, P_z) is projected on the sensor chip with distance P'_y from the center axis of projection.

We can now solve the problem of projection by using simple formula, revealed in the Eq. 6.7. If we continue with the transformation, we can write the result of Eq. 6.7 in the pixel coordinate system (as H_y) in Eq. 6.8

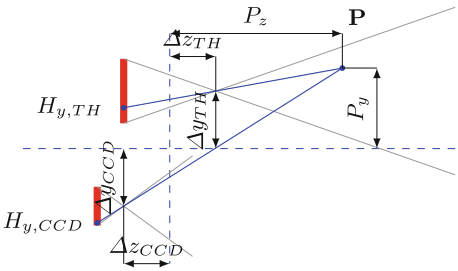

Fig. 6.4 To calibration of two cameras

$$P'_x = S \cdot \frac{P_x}{P_z}, \quad P'_y = S \cdot \frac{P_y}{P_z} \tag{6.7}$$

$$H_x = R_x\left(\frac{1}{2} - \frac{P'_x}{L_x}\right), \quad H_y = R_y\left(\frac{1}{2} - \frac{P'_y}{L_y}\right) \tag{6.8}$$

After combining Eqs. 6.7 and 6.8, we can write final projection transformation between original axis system (with center in the focal spot of the camera optics) and camera pixel system (with top-left center zero) in vector Eq. 6.9.

$$\mathbf{H} = 0.5 \cdot \mathbf{R} - \left(\frac{R_x \cdot S \cdot P_x}{L_x \cdot P_z}, \frac{R_y \cdot S \cdot P_y}{L_y \cdot P_z}\right) \tag{6.9}$$

$$\mathbf{P}_{TH} = \mathbf{P} - \Delta_{TH}, \quad \mathbf{P}_{CCD} = \mathbf{P} - \Delta_{CCD} \tag{6.10}$$

In the next step, we have two cameras, and we can parametrise their alignment in fixture as Δ, by distance of focus point to the center of the fixture in all axes parallell with camera's axes. The $\Delta = (\Delta_x, \Delta_y, \Delta_z)$ can be assumed in Eq. 6.9 to apply this shifts. The pixel transforms will be the same as in Eq. 6.11.

The final equation of transformation of point P to pixel index H is shown in Eq. 6.11.

$$\mathbf{H} = 0.5 \cdot \mathbf{R} - \left(\frac{R_x \cdot S \cdot (P_x - \Delta_x)}{L_x \cdot (P_z \Delta_z)}, \frac{R_y \cdot S \cdot (P_y - \Delta_y)}{L_y \cdot (P_z - \Delta_z)}\right) \tag{6.11}$$

Let's assume that we can control the calibration process, and we can trim the offset of focus points of cameras in z axis to zero ($\Delta_z \to 0$). The Eq. 6.11 will be slightly optimized to resulting Eq. 6.12b. The $\mathbf{A}$ vector is constant, and contains only parameters of the sensor. The value of $\mathbf{A}$ is shown in Eq. 6.16.

$$\mathbf{H} = 0.5 \cdot \mathbf{R} - \left(\frac{R_x \cdot S \cdot (P_x - \Delta_x)}{L_x \cdot P_z}, \frac{R_y \cdot S \cdot (P_y - \Delta_y)}{L_y \cdot P_z}\right) \tag{6.12a}$$

$$\mathbf{H} = 0.5 \cdot \mathbf{R} - \frac{1}{P_z}(A_x P_x, A_y P_y) + \frac{1}{P_z}(A_x \Delta_x, A_y \Delta_y) \tag{6.12b}$$

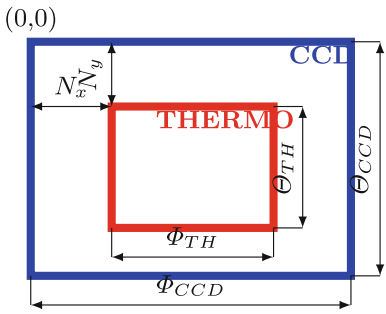

Fig. 6.5 Aligning thermo picture inside CCD picture

$$\mathbf{A} = \left(\frac{R_x \cdot S}{L_x}, \frac{R_y \cdot S}{L_y}\right) \tag{6.13}$$

Now we will assume that we are like to display both informations in one display. This can be seen on Fig. 6.5. Let the screen resolution of the monitor is $R_{x,SCR}$, $R_{y,SCR}$, and assume that every pixel is rectangular. We can map the camera signal with bigger field of view onto screen, like as stretched. The image of camera view will be placed at point (0, 0) and if the assumption, that $R_{x,SCR}$, $R_{y,SCR}$ is made, it will be stretched with coefficient K_{CCD} shown on Eq. 6.14a.

$$K_{CCD} = \frac{R_{x,SCR}}{R_{x,CCD}} \tag{6.14a}$$

$$K_{TH} = \frac{R_{x,SCR}}{R_{x,CCD}} \cdot \frac{\Phi_{TH}}{\Phi_{CCD}} \tag{6.14b}$$

After that, the thermometer signal will be mixed with CCD data in the same way, but the stretching coefficient will be slightly more complicated, and is solved in Eq. 6.14b.

The offset of thermocamera signal in the screen space can be explained mathematically from 6.15a and 6.15b.

$$N_x = \frac{1}{P_z} \cdot \frac{R_{x,SCR}}{R_{x,CCD}} \cdot \frac{1}{\Phi_{CCD}} \cdot \left(A_{x,CCD} \cdot \Delta_{x,CCD} \cdot \Phi_{CCD} - A_{x,TH} \cdot \Delta_{x,TH} \cdot \Phi_{TH}\right) \tag{6.15a}$$

$$N_y = \frac{1}{P_z} \cdot \frac{R_{x,SCR}}{R_{x,CCD}} \cdot \frac{1}{\Phi_{CCD}} \cdot \left(A_{y,CCD} \cdot \Delta_{y,CCD} \cdot \Phi_{CCD} - A_{y,TH} \cdot \Delta_{y,TH} \cdot \Phi_{TH}\right) \tag{6.15b}$$

For clarity, the Eqs. 6.15a and 6.15b can be simplified to form 6.16, where **Q** is constant.

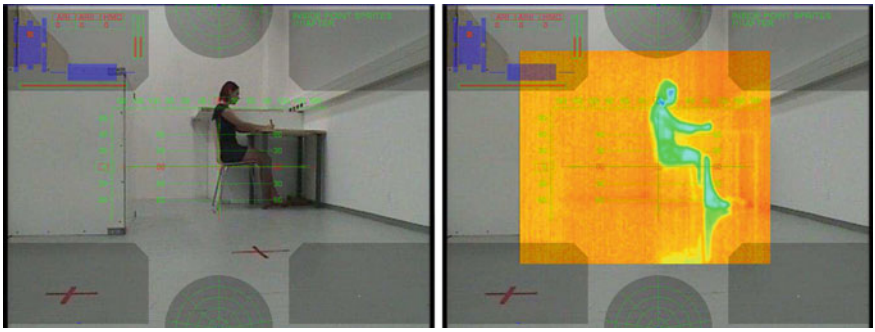

Fig. 6.6 CCD image (*left*), corresponding thermo-image placed over it (*right*)

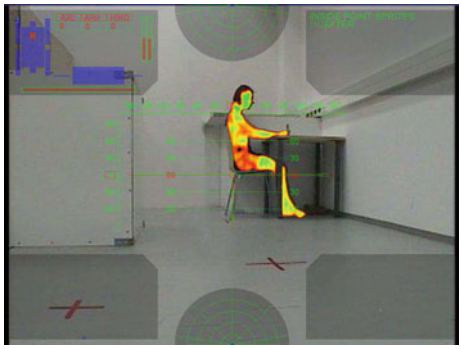

Fig. 6.7 Visible spectrum and thermovision data fusion

$$\mathbf{N} = \frac{1}{P_z}\left(Q_x, Q_y\right) = \frac{1}{P_z}\mathbf{Q} \tag{6.16}$$

The example of the described technique can be seen on Figs. 6.6 and 6.7. These figures are screenshots from Cassandra. The whole thermovision layer (i.e. RGB-to-HSV conversion, temperature filtration) is done by pixel-shaders in .NET, so the calculations are done on graphical card [15, 16].

To conclude this subsection, we can say there is almost a philosophical question how to display the frequencies of electromagnetic spectra that humans cannot see. The problem is much more complex in our case, where the image made by these imperceptible colors is mixed with the one we know (visible spectra). The technique described heretofore takes advantage of the imperfection of CCD imagers and commonly used displays together with the ability of our brain to adjust to something that does not represent reality perfectly, and to more clearly distinguish something uncommon.

6.4 Conclusions

The CCD camera and thermovision camera data-fusion has been commonly used on Orpheus-Explorer, and was demonstrated e.g. on IDET 2011 world exhibition. Although the user experience is very good, our team already works on more sophisticated method that includes CCD camera stereovision pair, thermal imager pair and TOF camera.

Acknowledgments This work was supported by the project CEITEC—Central European Institute of Technology (CZ.1.05/1.1.00/02.0068) from European Regional Development Fund.

References

1. Jacoff A, Weiss B, Messina E (2003) Evolution of a performance metric for urban search and rescue robots. In: Performance metrics for intelligent systems workshop, Aug 2003. Gaithersburg, MD
2. Zalud L (2004) Rescue robot league—1st place award winner. In: RoboCup 2003: robot soccer world cup VII. Springer, Germany. ISBN 3-540-22443-2
3. Wise E (1999) Applied robotics. Prompt Publications, USA. ISBN 0-7906-1184-8
4. Zalud L (2001) Universal autonomous and telepresence mobile robot navigation. In: 32nd international symposium on robotics. ISR 2001, pp 1010–1015, Seoul, Korea
5. Zalud L (2005) ORPHEUS reconniaissance teleoperated robotic system, In: 16th IFAC world congress, pp 1–6, Prague, Czech Republic
6. Martin CM, Moravec HP (1996) Robot evidence grids. The Robotics Institute Carnegie Melon University, Pittsburgh, 15213
7. Mullet K, Sano D (1995) Designing visual interfaces communication oriented techniques. Sun Microsystems Inc, USA. ISBN 0-13-303389-9
8. Oyama E, Tsunemoto N, Tachi S, Inoue S (1993) Experimental study on remote manipulation using virtual reality. Presence 2(2):112–124
9. Sheridan TB (1992) Telerobotics, automation, and human supervisory control. MIT Press, Cambridge
10. Ayache N (1991) Artificial vision for mobile robots stereo vision and multisensory perception (translation). The MIT Press, Cambridge. ISBN 0-262-01124-7
11. Gonzalez G, Woods RE (2002) Digital image processing, 2nd edn. Prentice Hall, Englewood Cliffs. ISBN 0-201-18075-8
12. Everett HR (1995) Sensors for mobile robots, theory and applications. AK Peters Ltd, USA. ISBN 1-56881-048-2
13. FlirSystems (2007) Retrieved Mar 12 2007. from http://www.flirthermography.com
14. Wyszecki G, Stiles WS (2000) Color science concepts and methods, quantitative data and formulae. Wiley-Interscience, New York. ISBN 0-471-02106-7
15. LaMothe A (2003) Tricks of the 3D game programming gurus advanced 3D graphics and rasterization. SAMS Publishing, USA. ISBN 0-672-31835-0
16. Luna DF (2003) Introduction to 3D game programming with DirectX 9.0. Wordware Publishing Inc, USA. ISBN 1-55622-913-5