

# Thermalkamera zur Stresserkennung im Creative Space

Jorin Kleimann

Hochschule für angewandte Wissenschaften Hamburg  
Fakultät Technik und Informatik, Department Informatik  
Berliner Tor 7, 20099 Hamburg  
[jorin.kleimann@haw-hamburg.de](mailto:jorin.kleimann@haw-hamburg.de)  
[www.haw-hamburg.de](http://www.haw-hamburg.de)

**Abstract.** In der folgenden Ausarbeitung werden durchgeführte Versuche, Technologien und zukünftige Projektschritte zur Implementierung einer Thermalkamera in den Creative Space und das EmotionBike der HAW Hamburg erläutert und diskutiert. Außerdem wird ein Blick auf den aktuellen Stand der Forschung in diesem Bereich geworfen und ein Zukunftsausblick auf die Masterthesis gewährt.

**Keywords:** computer vision, human-computer-interaction, thermal imaging, light field

## 1 Einleitung

### 1.1 Motivation

Es gibt neben herkömmlichen Kameras, die in der Lage sind Normalbilder aufzunehmen, eine Reihe an weiteren modernen Kameratechnologien, deren Bilder weitaus mehr Informationen enthalten als nur unterschiedlich eingefärbte Pixel. Lichtfeldkameras, die durch eine große Anzahl an sogenannten Mikrolinsen die zusätzlich zur Hauptlinse ein Objekt aus verschiedensten Winkeln aufnehmen, besitzen die Möglichkeit nachträglicher Fokussierung einzelner Bereiche eines Bildes. Thermalkameras, gerade jene welche in einem Temperaturbereich zwischen 0,03K und 0,05K arbeiten und somit schon zu den hochauflösenden Thermalkameras zählen, werden heute in Industrie, Handwerk und Forschung gleichermaßen eingesetzt. Feinste Temperaturunterschiede lassen sich in einem gewissen Zeitverlauf messen und so die Ergebnisse analysieren.

Diese beiden genannten Technologien birgen für den Bereich Computer-Vision bzw. Bildverstehen großes Potenzial. Sowohl im Bereich der Gestenerkennung, als auch im Erkennen von Objekten oder beispielsweise Gesichtern. Mit Hilfe einer Lichtfeldkamera könnten sich Gesten oder Gesichtsausdrücke erkennen lassen, mit einer Thermalkamera beispielsweise Temperaturveränderungen.

Der Anreiz mit solchen hochmodernen Technologien zu arbeiten war unter anderem ausschlaggebend für die Auseinandersetzung mit dem Thema meinerseits. Aufgrund der recht hohen Kosten für ebendiese Technologien ist die Forschung an dieser Stelle auch noch nicht am Ende.

## 1.2 Arbeitsgruppe

Angesiedelt ist die in dieser Arbeit beschriebene Projektarbeit und Forschung sowohl im Living Place als auch im Creative Space von Kai von Luck und Susanne Draheim. Diese beiden Arbeitsumgebungen fallen in die Bereiche Human-Computer-Interaction, smart home und virtual bzw. augmented reality und werden im folgenden näher erläutert.

**Living Place** Der Living-Place befindet sich auf dem Campus der HAW am Berliner Tor. Er deckt verschiedene Bereiche des IT-unterstützten urbanen Lebens ab. Die 140 Quadratmeter große Wohnung mit Küche, Wohnbereich und Schlafzimmer ist durch seinen Aufbau als echtes Apartment ideal für Experimente [3]. Ebenfalls im Living Place angesiedelt ist das sogenannte EmotionBike. Hierbei handelt es sich um ein feststehendes Fahrrad auf der Basis eines Ergometers auf dem ein Proband durch eine virtuelle Welt fahren kann, die einem auf einem Bildschirm vor dem Fahrrad dargestellt wird. Dabei erfassen Kameras und Sensorik die Mimik und physiologische Daten des Probanden um während der Fahrt emotionale Zustände wie Stress zu erkennen. Das Ziel dabei ist es nicht nur die Emotionszustände zu erkennen, sondern in der Lage zu sein, auf diese zu reagieren. Dieser Ansatz nennt sich *companion*.

**Creative Space** Der *Creative Space for Technical Innovations* wird seit Oktober 2015 an der HAW in der Fakultät für Technik und Informatik aufgebaut. Der Schwerpunkt liegt auch hier im Bereich der Human-Computer-Interaction. Finanziert wird der Creative Space, welcher sich im Steindamm unweit des HAW Campus befindet, aus dem Zukunftsfond der HAW Hamburg und anteilig von der Hamburger Behörde für Wirtschaft, Verkehr und Innovation. Die Eröffnung ist für das Sommersemester 2016 geplant [1]. Der für diese Projektarbeit wichtige Bereich des Creative Space ist die Augmented-Reality und Virtual-Reality Umgebung. Virtuelle Realität beschreibt das Wahrnehmen einer virtuellen Umgebung in einer in Echtzeit computergenerierten Welt[5].

## 2 Grundprojekt

Im Grundprojekt wurden erste technologische Schritte sowohl mit Lichtfeldkamera als auch mit Thermalkamera gemacht. Ziel dabei war es, die Technologien kennenzulernen und ihre etwaigen Grenzen auszumachen um festzustellen ob eine weitere Anwendung mit der Thermalkamera als auch mit der Lichtfeldkamera realisierbar ist. Dieses Kapitel ist in drei Unterpunkte unterteilt. Im ersten Teil wird das Verfahren mit der Lichtfeldkamera erläutert und geklärt aus welchem Grund diese Kamera für weitere Projektschritte ausgeschlossen wurde. Im zweiten und dritten Teil wird näher auf die Thermalkamera und die entwickelte Anwendung eingegangen. Außerdem werden die Grundlagen der Thermografie erläutert und ein Vorversuch zur Stresserkennung im Thermalbild vorgestellt.

## 2.1 Lichtfeldkamera

Die Funktion und Technologie die in der Lichtfeldkamera R5 der Firma Raytrix steckt wurde in meinem Paper im Grundprojekt umgehend erläutert und wird deshalb hier nicht weiter angeführt. In einem ersten Versuch wurden die Grenzen und Möglichkeiten der Lichtfeldkamera erforscht um folgende Fragen zu klären:

- Ist eine saubere Kalibrierung der Kamera möglich?
- Sind die qualitativen Ergebnisse der Bilder zufriedenstellend?
- Inwiefern beeinflussen äußere Umstände (Natürliches Licht/Tageslicht) die Bilder?

Um diese Fragen beantworten zu können wurde eine kontrollierte Umgebung eingerichtet mit einer Szene in der verschiedene Tiefen und Objekte zu sehen waren.



Fig. 1: Beispiel der Firma Raytrix

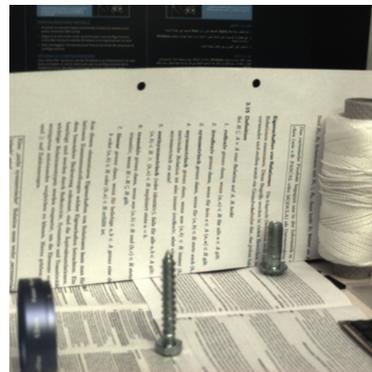


Fig. 2: Eigenes Bild mit Raytrix R5

In Figur 1 ist das "Paradebeispiel" der Firma Raytrix abgebildet, welches aus den Werbebildern zur Raytrix R5 entnommen wurde und dementsprechend erwartet werden kann, dass es mit einer Raytrix R5 Lichtfeldkamera aufgenommen wurde. Auf diesem Bild ist zu erkennen, dass alle Bereiche des Bildes fokussiert sind, obwohl verschiedene Tiefen vorhanden sind, also genau das, was man mit einer Lichtfeldkamera erreichen möchte. In Figur 2 ist die Versuchsszene die unser Raytrix R5 aufgenommen wurde. Hier wird auf den ersten Blick deutlich, dass nicht alle Bereiche des Bildes im Fokus sind (zu erkennen an der Schrift der schwarzen Linsenabdeckung vorne links). Auch ein nachträgliches Fokussieren bringt hierbei nicht das gewünschte Ergebnis. Denn um im späteren Verlauf die exakten Tiefeninformationen zu bekommen, ist ein Fokussieren von allen Bereichen eines Bildes von Nöten. Wenn man sich nun die Tiefeninformationen von Figur 2 anschauen würde, erhält man ein enormes Rauschen, welches sich in einer Art "wabern" beziehungsweise "flackern" zeigt, was unter realen Bedingungen zu keinen dauerhaft zuverlässigen Informationen führt.

Nachdem dieses für das Vorhaben nicht zufriedenstellende Ergebnis unter "perfekten" Bedingungen zustande gekommen ist und die Firma Raytrix jedoch nach einer Rücksprache das Ergebnis als Zufriedenstellend bezeichnet hat, wurde der Beschluss gefasst die Lichtfeldkamera als Technologie für das weitere Projekt vorerst auszuschließen.

## 2.2 Applikation Thermalkamera

Nach einem vorzeitigen Ausschluss der Lichtfeldkamera für das weitere Projekt wurde das Augenmerk auf die Thermalkamera gelegt. Die von der HAW Hamburg erworbene Thermalkamera ist das Modell VarioCam HD der Firma Infratec. Diese besitzt eine Temperaturauflösung von 0,03K bei 30 °C, welche im Vergleich mit anderen Herstellern sehr weit vorne liegt. Die Thermalkamera ist in der Lage thermische Strahlung im Spektralbereich von 7,5-14  $\mu\text{m}$  zu messen. Das Bildformat im Thermalbild (sprich Infrarotpixel) ist 1024x768.

**Grundlagen der Thermografie** Jeder Körper mit einer Temperatur oberhalb des absoluten Nullpunkts (0K,  $-273,15\text{ °C}$ ) sendet elektromagnetische Strahlung aus. Dabei befindet sich die Infrarotstrahlung auf einer Wellenlänge zwischen 1  $\mu\text{m}$  und 100  $\mu\text{m}$ . Zur Messung der Infrarotstrahlung eines Körpers wird die von ihm ausgestrahlte Energie durch eine spezielle Linse auf einen Infrarotdetektor innerhalb der Thermalkamera fokussiert. Dabei übernimmt dann die Sensorelektronik die Bildverarbeitung. Nach der Bestimmung der Intensität dieser Strahlung wird anschließend daraus die Temperatur des aussendenden Körpers ermittelt [2].

Für adäquate Infrarotmessungen ist die Beschaffenheit der Messstrecke von hoher Bedeutung. Dabei können nur Körper und Objekte gemessen werden die nicht von anderen stark strahlenden Objekten verdeckt werden. Sonneneinstrahlung und weitere sogenannten Störstrahlung müssen dabei ebenfalls rausgerechnet werden.

Das der Thermalkamera von Infratec beiliegende SDK ist proprietär für die Sprachen C++ und Delphi entworfen. Um für die weiteren Versuche und ein späteres Facetracking im Thermalbild wurde eine erste Anwendung geschrieben, die in der Lage ist Einzelbilder und Sequenzen von der Kamera aufzunehmen. Hierzu wurde ein C#-Interface entwickelt, welches die Funktionen der Infratec DLL implementiert. Die Thermalkamera ist über ein herkömmliches RJ45-Kabel mit einem Computer verbunden, auf dem die Anwendung läuft. Der hauptsächliche Grund für das Entwickeln einer C#-Anwendung war, dass die spätere Verarbeitung des Datenstreams von der Kamera ebenfalls in C# geschehen soll. Diese C#-Applikation dient im folgenden Hauptprojekt und der Masterarbeit dann die Grundlage für das weitere Vorgehen, bzw. diese Applikation wird dann um weitere Komponenten erweitert.

### 2.3 Stresserkennung im Thermalbild

Für die Erkennung von Stress im Gesicht mit Hilfe des Thermalbilds sind im Grunde zwei Bereiche des menschlichen Gesichts wichtig: die Stirn und der Wangenbereiche. Dabei gilt die Annahme, dass sich die Erhöhung der Gesichtstemperatur aufgrund der Wärmeübertragung des Blutflusses an der Hautoberfläche mit einer Thermalkamera messen lässt.

Zur Verifikation dieser Ansätze ist ein Versuch gestartet worden, der verschiedene Gesichter in mit Absicht hervorgerufenen Stresssituationen aufgenommen hat. Hierzu wurde ein Pecha-Kucha ähnlicher Vortrag entworfen, der von Probanden vorgetragen werden musste.

Pecha-Kucha ist eine ursprünglich aus dem japanischen stammende Vortragsweise, bei der ein Vortrag aus genau 20 Bildern besteht, die jeweils genau 20 Sekunden gezeigt werden und die Bilder automatisch wechseln. Das heißt, dass der Vortragende in keiner Weise einen Einfluss darauf hat, wann das nächste Bild gezeigt wird. In der originalen Pecha-Kucha Variante hat jeder Vortragende seine eigenen Bilder, die er zeigen möchte. In der für diesen Versuch gewählten Variante haben die Mitglieder des Projektteams die Bilder vorbereitet. Das Thema für diesen Versuch lautete "Hamburg" was allen Probanden einen eventuell ähnlichen Kenntnisstand ermöglicht und so die Voraussetzungen für jeden Probanden gleich sind. Während des Vortrags wurde alle 5 Sekunden ein Bild mit der Thermalkamera aufgenommen. Im Nachhinein werden dann die Temperaturunterschiede an der Stirn ausgewertet. Die folgende Abbildung zeigt die Erwärmung eines Stirnbereiches in einer Stresssituation während des Vortrages. Mit Hilfe dieser Messungen war die Anfangs vermutete These, dass sich

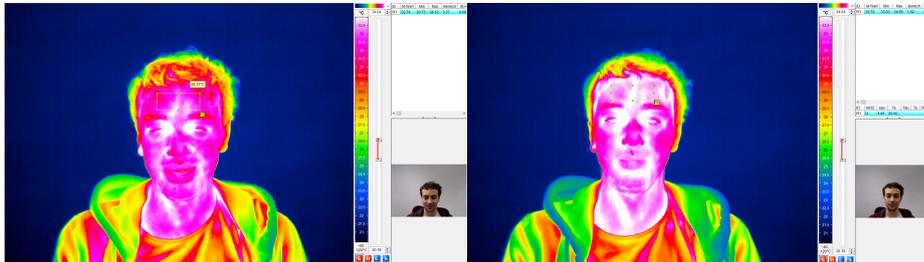


Fig. 3: Durchschnittliche Temperatur vorFig. 4: Durchschnittliche Temperatur  
der Stresssituation: 34,12 °C während der Stresssituation: 34,55 °C

bestimmte Bereiche des Gesichts in einer bestimmten Situation, beispielsweise bei Stress, erwärmen, bestätigt.

## 3 Hauptprojekt

Das Hauptprojekt ist ebenfalls in drei Schritte unterteilt. Im ersten Teil wird die Thermalkamera in die Umgebung des EmotionBike implementiert, so dass

der dort verwandte MessageBroker die Thermalkamera mit der anderen Sensorik synchronisieren kann und Aufnahmen gestartet werden können. So kann ein Proband auf dem Fahrrad zusätzlich mit der Thermalkamera aufgenommen werden und später sein Emotionslevel analysiert werden. Im zweiten Teil soll die Analyse, die bis hierher noch per Hand geschieht (Messbereich im Gesicht auswählen und Temperatur "rausziehen") automatisiert werden. Der erste Schritt hierzu ist das Tracken des Gesichts vom Probanden. Denn um immer an der selben Stelle im Bild messen zu können, muss das Gesicht vom Rest des Bildes extrahiert werden können. Sobald dies Möglich ist, soll im dritten Schritt eine Implementierung in den Creative Space geschehen, um dann auch bei Gesichtern die eventuell abgedeckt sind, sich viel Bewegen oder in verschiedene Richtungen schauen, messen zu können.

### 3.1 Versuch im Emobike

Um noch mehr Daten aufnehmen zu können wurde die Thermalkamera bzw. die C#-Applikation in das Emotionbike integriert. Hierzu wurde der MessageBroker und die dazugehörigen Messages implementiert. Gegen Ende des Sommersemesters 2016 startet eine größere Versuchsreihe am Emotionbike bei der ca. 20 Probanden auf dem Fahrrad sitzen werden. Durch die MessageQueue wird die Thermalkamera diese Versuche komplett mit aufzeichnen. Mit diesem so entstehenden Datensatz sind dann genug Thermalbild-Daten vorhanden um im späteren Verlauf auf diesen zu arbeiten. Durch die Synchronisierung über die Message Queue ist dann auch eine eventuell später notwendige Kalibrierung mit den anderen Kameras am EmotionBike (Kinect, Eyetracker) möglich.

### 3.2 Facetracking im Thermalbild

Wie Eingangs erwähnt wird für eine automatisierte Analyse das Tracken des Gesichts eines Probanden benötigt. Durch Bibliotheken wie OpenCV [4] ist ein Tracking des menschlichen Gesichts im Normalbild schon jetzt ohne große Hürden nutzbar. Da ein Thermalbild jedoch in der Regel weniger Kanten als ein Normalbild liefert, funktioniert dieses Tracking noch nicht zufriedenstellend.

Mit Hilfe der im Versuch im EmotioBike entstandenen Bilder soll deshalb OpenCV für ein Tracking von Thermalbildern angelernt werden. Dies geschieht mit Hilfe sogenannter haar-ähnlicher Merkmale. Hierbei wird der in einem Bild gesuchte Bildausschnitt als ein Vektor mit bestimmten (haar-ähnlichen) Merkmalen repräsentiert. Dieser Vektor wird über das Bereitstellen einer großen Anzahl an positiven Bildern, also Bilder die einen gesuchten Bildausschnitt (hier: das menschliche Gesicht als Thermalbild) enthalten, definiert. Ist das Anlernen abgeschlossen wird ein Bild anhand der Merkmale über ein stumpfes Übereinanderlegen von Bildausschnitt und Bild nach eben diesen Merkmalen durchsucht.

### 3.3 Implementierung in Creative Space

Sobald ein Tracking zuverlässig funktioniert ist es möglich, dass die Analyse von Stress beziehungsweise auch anderen Emotionen auch im Creative Space in augmented oder virtual-reality Anwendungen stattfindet. So wäre ein denkbarer Versuch, die Frage zu beantworten, ob sich die sogenannte *motion sickness* auch mit Hilfe der Thermalkamera erkennen lässt. Die *motion sickness* oder auch *Simulatorkrankheit* tritt dann ein, wenn dem visuellen Sinn eine Bewegung vorgetäuscht wird, die der vom realen Orientierungs und Lagesinn entsprechenden Informationen nicht entspricht. Dies kann zu Übelkeit, erhöhtem Speichelfluss, Benommenheit, Schwindelgefühlen und sogar Erbrechen führen [5]. Wenn es möglich ist diesen Gemütszustand zu erfassen, kann man einen Nutzer vor dem Eintreten der motion sickness warnen und so früh genug eine virtuelle Anwendung abbrechen.

Ein weiterer Einsatzzweck im Creative Space wäre die Unterstützung für augmented-reality Anwendungen. Mit Hilfe der Thermalkamera ist es möglich über die hinterlassenen Wärmeflecken einer Hand oder eines Fingers auf einem Objekt zu erkennen, ob ein Objekt berührt wurde oder nicht. Diese Idee ist im Kapitel 4 in der Erläuterung des Papers *Thermal Touch: Thermography-Enabled Everywhere Touch Interfaces for Mobile Augmented Reality Applications* näher betrachtet.



Fig. 5: Hand bei Berühren eines Tisches

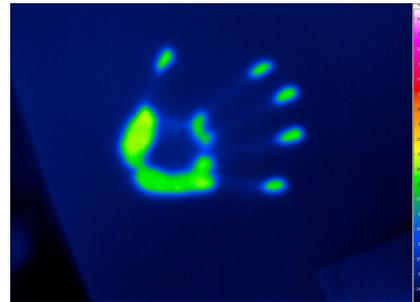


Fig. 6: Thermalabdruck auf dem Tisch

## 4 Aktueller Stand der Forschung

Im folgenden Abschnitt sind drei ausgewählte Arbeiten aufgeführt, die ähnliche Ansätze verfolgt haben wie sie auf den letzten Seiten vorgestellt wurden.

**Thermografische Erfassung von Stress in simulierten Vorstellungsgesprächen [6]** In dieser Arbeit wurde zur Erfassung von Stress ein ähnlicher Ansatz gewählt. Wie der Titel schon sagt, hat das Forschungsteam in simulierten

Vorstellungsgesprächen versucht Stress zu erzeugen und diesen mit Hilfe einer Thermalkamera zu erfassen. Die Probanden mussten vor "Sprachexperten" eine 5-minütige Rede auf Englisch halten. Hierfür wurde eine Vorbereitungszeit von 10 Minuten vorgegeben. Die Ergebnisse waren ähnlich derer, die im hiesigen Versuch herauskamen: Stress kann Wärme in gewissen Bereichen des menschlichen Gesichts hervorrufen.

**Vision-Based pedestrian detection using haar-like features [8]** In dieser Arbeit werden haar-ähnliche Merkmale benutzt um Fußgänger als Objekt in Normalbildern anzulernen. Die Anwendung findet auch hier in einem Echtzeitsystem statt, welches einen Videostream direkt absucht und in dem recht gute Bilderkennungsraten erzielt werden.

**Thermal Touch: Thermography-Enabled Everywhere Touch Interfaces for Mobile Augmented Reality Applications [7]** Dieses Paper beschreibt einen Ansatz, der auch in der Medienwelt für ein wenig aufsehen gesorgt hat. Berührt man mit seinen Händen ein Objekt in einem Raum so hinterlässt dieses Berühren einen "Wärmefleck" auf diesem Objekt (es sei denn dieses Objekt ist über 37°C warm). Schaut man sich also dieses Objekt an, so erkennt man den Finger bzw. Handabdruck nach der Berührung auf dem jeweiligen Objekt für einige Sekunden. Dieses "Phänomen" wurde dann dazu genutzt in einer augmented-reality Umgebung das Berühren eines Objektes zu erkennen. Dabei wurde eine Thermalkamera mit einer normalen Kamera kalibriert. Das Objekt an sich wurde mit der normalen Kamera getrackt und das Berühren des Objektes dann mit Hilfe der Thermalkamera. Eine Nutzung dieses Phänomens wäre beispielsweise in ähnlicher Form für den Creative Space denkbar.

## 5 Ausblick

Einige der Ziele für die Masterarbeit wurden in den vorangegangenen Kapiteln bereits genannt. Allerdings wird in der Masterarbeit neben der Fertigstellung der Anwendung für die automatisierte Analyse der Thermalbilder auch der wissenschaftliche Bereich der Analyse abgehandelt werden. Dabei muss dann untersucht werden, inwiefern die Temperaturdaten und deren Schwankungen überhaupt zuverlässig und aussagekräftig sind. Die Frage, was bestimmte Informationen aus den Thermalbildern bedeuten, verlangen die Auseinandersetzung mit physiologischen und biologischen Grundlagen.

Zusammenfassend ist zu sagen, dass in dieser Ausarbeitung neben der Vorstellung der Arbeitsgruppe, technologischen Grundlagen vor allem die bereits erledigte Projektarbeit vorgestellt wurde. Außerdem wurde ein Überblick darüber gewährt, in welche Richtung die Arbeit im Hauptprojekt - über das Implementieren des Gesichtstrackings im Thermalbild bis hin zu einer Analyse der Temperaturschwankungen im Gesicht - gehen wird. Des Weiteren wurde ein Einblick in verwandte Arbeiten und ein Ausblick auf die Masterthesis gewährt.

## References

1. Creative Space Hamburg. <http://creative-space.haw-hamburg.de/>
2. Einführung in die Theorie und Praxis der Infrarot-Thermografie. [www.thermografie.co.at/files/infratec.pdf](http://www.thermografie.co.at/files/infratec.pdf)
3. Living Place Hamburg. [http://livingplace.informatik.haw-hamburg.de/blog/?page\\_id=47](http://livingplace.informatik.haw-hamburg.de/blog/?page_id=47)
4. Open Source Computer Vision. <http://opencv.org/>
5. Dörner, R., Broll, W., Grimm, P., Jung, B.: Virtual und Augmented Reality (VR / AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität. eXamen.press, Springer Berlin Heidelberg (2014), <https://books.google.de/books?id=aXckBAAAQBAJ>
6. Krajewski, J., Schnieder, S., Datkiewicz, S.: Thermografische Erfassung von Stress in simulierten Vorstellungsgesprächen. In: 48. Kongress der Deutschen Gesellschaft für Psychologie (2012)
7. Kurz, D.: Thermal touch: Thermography-enabled everywhere touch interfaces for mobile augmented reality applications. In: Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on. pp. 9–16 (Sept 2014)
8. Monteiro, G., Peixoto, P., Nunes, U.: Vision-based pedestrian detection using haar-like features. *Robotica* 24, 46–50 (2006)