

Höher, besser, schneller

Videoanalyse von Sportspielen

(Ein Beitrag für den Deutschen Studienpreis 2002 von Anna Egorova)



Anna Egorova, Alexander Gloye, Raúl Rojas

Technical Report B-15-2002

November 2002

Freie Universität Berlin

Institut für Informatik

Takustr. 9

14195 Berlin

{egorova, gloye, rojas}@inf.fu-berlin.de

Höher, besser, schneller: Videoanalyse von Sportspielen

(Ein Beitrag für den Deutschen Studienpreis 2002 von Anna Egorova)

Anna Egorova, Alexander Glove, Raúl Rojas
Technical Report B-15-2002, November 2002

Freie Universität Berlin
Institut für Informatik
Takustr. 9
14195 Berlin
{egorova, glove, rojas}@inf.fu-berlin.de

Zusammenfassung

„Höher, besser, schneller“ lautete das Motto der Olympischen Spiele, als sie 776 v.Ch. zum ersten Mal durchgeführt wurden. Heutzutage sieht es so aus, als ob sich alles in unserem Leben um diese drei Wörter dreht.

Die Eisenbahn wird schneller, das Internet wird immer grösser, die Computer werden immer kleiner und verarbeiten immer mehr Informationen, die Software-Entwickler bieten immer schnellere, bessere und bequemere Programme, die Unterhaltungsmedien werden immer interessanter und realistischer und die Forschungseinrichtungen und die Universitäten versuchen immer wieder, bessere, schnellere und einfachere Lösungen für alles anzubieten.

Die Analyse von Sportspielen sind für alle interessant – egal, ob ein Fussballfan spät abends wissen möchte, wie oft der Torwart seiner Lieblingsmannschaft den Ball abgewehrt hat; oder ob ein Trainer seine Strategie prüfen will; oder ob Robotik-Forscher versuchen, ihren Robotern Fussball beizubringen. Die vollständige Analyse durch den Menschen erfordert allerdings sehr viel Zeit sowie Kenntnisse und Zeit hat man heutzutage einfach nicht. Die einzige Möglichkeit, alle Informationen aus einem Fussball-Spiel zu gewinnen, ist diese Aufgabe einem Computer zu überlassen.

Das vorgestellte System „CatchIt!“ basiert auf verschiedenen modernen Methoden der Bildverarbeitung und bietet die Möglichkeit, ein aufgenommenes Fussballspiel zu analysieren und die wichtigen Informationen herauszufiltern. Das System verfolgt die einzelnen Spieler und den Ball, liefert Statistiken über die gelaufenen Kilometer, über die Anzahl der Ballkontakte, über die in der eigenen Hälfte verbrachte Zeit usw.

An der Freien Universität Berlin werden seit 1997 Fussballroboter gebaut. Im Rahmen des Forschungsprojektes „RoboCup“ wird auch das Video-Analyse-System benutzt, um Strategien herauszufinden und Statistiken zu erstellen.

1. Einführung

Fussball wurde in England erfunden und auf den Strassen von London mit einem kleinem Stoffball gespielt. Es war nur ein Spiel, wo jeder seinen Spass hatte und die Möglichkeit ausnutzte, mit Freunden und Bekannten die Zeit zu vertreiben.

Heutzutage sind Sportspiele und insbesondere das Fussballspiel ein sorgfältig geplantes und durchgeführtes Zusammenspiel von Strategie, Taktik und Können. Die Trainer und ihre Spieler verbringen Wochen und Monate vor wichtigen Spielen in Analyse von alten Spielen und Planen die Taktik für das bevorstehende Spiel. Diese Analyse beinhaltet Informationen wie die Positionen der Spieler, die in der eigenen Hälfte verbrachte Zeit, das Zusammenspiel von allen Mannschaftsmitgliedern, die gelaufenen Kilometer usw. Die einzige Möglichkeit, diese Analyse durchzuführen ist wieder und wieder die Aufnahmen von verschiedenen Spielen zu sehen und versuchen, die wichtigen Informationen zu gewinnen.

Die auf diese Weise gewonnen Informationen sind aber nicht nur für die Spieler und Trainer selbst interessant. Die Fans würden auch gern etwas mehr über ihre Lieblingsmannschaft erfahren, tiefer in die Organisation und Strategie blicken. Die Unterhaltungsmedien würden diese Informationen ihren Zuschauern gern anbieten (siehe **Abbildung 1**). Forscher würden gern mehr über das menschliche Verhalten wissen. Robotik-Entwickler würden gern dieselben Strategien bei ihren Fussball-Robotern anwenden.

Damit man eine Fussballspiel-Aufnahme zeitsparend und effizient analysieren kann, muss man, wie in fast allen anderen Bereichen unseres Lebens heutzutage, zu Computer und speziell entwickelten Systemen greifen, die das ganze Verfahren automatisch durchführen und die Ergebnisse speichern.

Das „CatchIt!“ – System wertet die Bilder aus einer oder mehreren statischen Videokameras aus. Die einzelnen Spieler und der Ball werden im Bild lokalisiert und ihre Koordinaten gespeichert. Durch die permanente Lokalisierung wird die Trajektorie für jeden einzelnen Spieler und für den Ball während des gesamten Spiels ermittelt. Dadurch erreicht man erstens eine sehr komprimierte Darstellung des Spiels, die nur die wichtigsten Informationen, wie die Positionen der Spieler und des Balls, beinhaltet. Beispielsweise braucht ein farbiges Bild, aufgenommen von einer Videokamera, mit 720x576 Pixel Auflösung (wie es bei den hier benutzten Testdaten der Fall ist) ungefähr 1,2 MB Speicherplatz. Für ein ganzes Spiel von 90 Minuten bedeutet das ein Speicherplatz von ungefähr 160 GB (etwa 250 CDs).



Abbildung 1. Beispiel für eine erstellte Statistik, wie sie z.B. im Fernsehen angewandt werden könnte.

Wenn man aber dagegen nur die aus der Analyse gewonnen Informationen speichert, die die Positionen der Spieler, der Schiedsrichter und des Balls enthalten, braucht man für jedes Objekt nur dessen Koordinaten zu speichern. Man speichert daher die Informationen für insgesamt 26 Objekte (22 Spieler, 3 Schiedsrichter und 1 Ball) in 52 Bytes. Für ein ganzes Spiel von 90 Minuten bedeutet das ein Speicherplatz von 0,3 MB. Man erreicht also eine Komprimierung von 1: 60 000!

Zweitens bekommt man eine vollständige Statistik über das Spiel. Darunter befinden sich Informationen wie:

- ✍ Die von jedem einzelnen Spieler gelaufenen Kilometer
- ✍ Spitzengeschwindigkeit eines Spielers
- ✍ Anzahl der Ballkontakte pro Spieler
- ✍ Ballkontrolle: mittlere Anzahl von Sekunden mit dem Ball vor der Abgabe
- ✍ Statistik der ganzen Mannschaft: gelaufene km, Ballkontakte, usw.
- ✍ Minuten verbracht in der eigenen Hälfte
- ✍ Minuten verbracht im Strafraum oder in der Nähe
- ✍ Die Formation der Verteidigung (Vierer- oder Dreierkette, usw.)
- ✍ Die Angriffsformation

Diese Daten bieten einerseits den Fernsehzuschauern die Möglichkeit, einen tieferen Einblick in die Strategie der Mannschaften zu werfen und liefern andererseits Trainern und Forschern wichtige Informationen und Statistiken.

2. Verteilung der Kameras

Die Verteilung der Videokameras bei der Aufnahme des Spiels spielt eine sehr grosse Rolle für die nachfolgende Analyse und für ihre Qualität und Genauigkeit. Die erste und die wichtigste Forderung ist dabei, mit Hilfe von einer oder mehrerer Kameras das ganze Spielfeld zu erfassen. Auf diese Weise können die Objekte im Bild automatisch verfolgt werden. In den hier durchgeführten Tests wurde das Bild mit einer Videokamera von der Tribüne, ungefähr in der Spielfeldmitte, aufgenommen (siehe **Abbildung 2**).

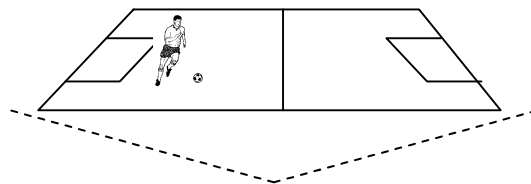


Abbildung 2. Bildaufnahme mit einer Videokamera aus der Tribüne

Im Fall von nur einer Videokamera entstehen aber mehrere Probleme. Da das ganze Spielfeld aufgenommen werden muss, werden die Objekte im Bild sehr klein. Ein Pixel (ein Punkt) im Bild entspricht so in den Testdaten z.B. einer Fläche von 15 x 15 cm. Man kann mit solchen ungenauen Daten also keine zuverlässigen Aussagen für die Identität der Spieler, ihren genauen Positionen, Orientierung, kleine Bewegungen usw. machen.

Ein anderes Problem stellt die Überdeckung von verschiedenen Objekten dar. Falls zwei Objekte hintereinander in der Sichtlinie der Kamera stehen, kann das System sie nicht mehr unterscheiden und verliert das hintere Objekt. Der Fall wurde auch in **Abbildung 5(a)** dargestellt.

Andererseits, wie man auch von den hier vorgestellten Ergebnissen sieht, gewinnt man die wichtigsten Informationen auch von Bilddaten, die nur von einer Videokamera aufgenommen wurden und die Aufzeichnung ist einfach.

Für die genaue Ermittlung der Trajektorien ist allerdings noch mindestens eine zweite Videokamera erforderlich, um die Bewegung der Objekte in allen Richtungen abschätzen zu können, wie es in **Abbildung 3** zu sehen ist.

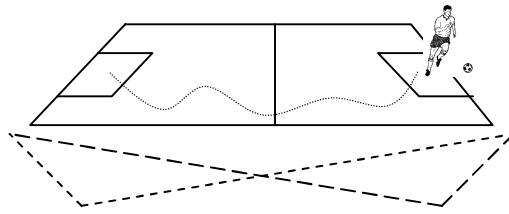


Abbildung 3. Bildaufnahme mit zwei Kameras aus der Tribüne

Dieser Fall hat verschiedene Vor- sowie Nachteile. Falls man die Kameras auf die beschriebenen Positionen bringt und aber von beiden Kameras das ganze Spielfeld umfasst, kann man die Überdeckungen der Spieler sehr gut umgehen (falls sich die Objekte vor der einen Kamera überdecken, werden sie sich mit Sicherheit vor der anderen nicht überdecken), aber die Grösse der Objekte bleibt im Vergleich zu der einzelnen Kamera unverändert.

Bei der zweiten Möglichkeit vertauschen sich die Vor- und Nachteile: man positioniert die beiden Kameras auf der Tribüne so, dass man von jeder eine Hälfte des Spielfeldes aufnehmen kann. Dann werden die Objekte zwar etwa doppelt so gross, aber das Problem der Überlappungen bleibt.

Die beste Lösung in diesem Fall wäre ein automatisches Kamera-System, das im Fall von Überdeckungen das ganze Spielfeld aufnehmen lässt und so das verlorene Objekt wieder findet.

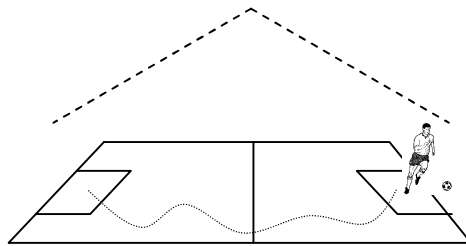


Abbildung 4. Bildaufnahme mit einer Kamera, positioniert über der Spielfeldmitte

Eine andere Möglichkeit ist, wenn man nur eine Kamera benutzt, die sich aber genau über die Spielfeldmitte befindet (**Abbildung 4**). In diesem Fall werden sich die Objekte nie überlappen und die richtigen Positionen liefern (**Abbildung 5 (c)**).

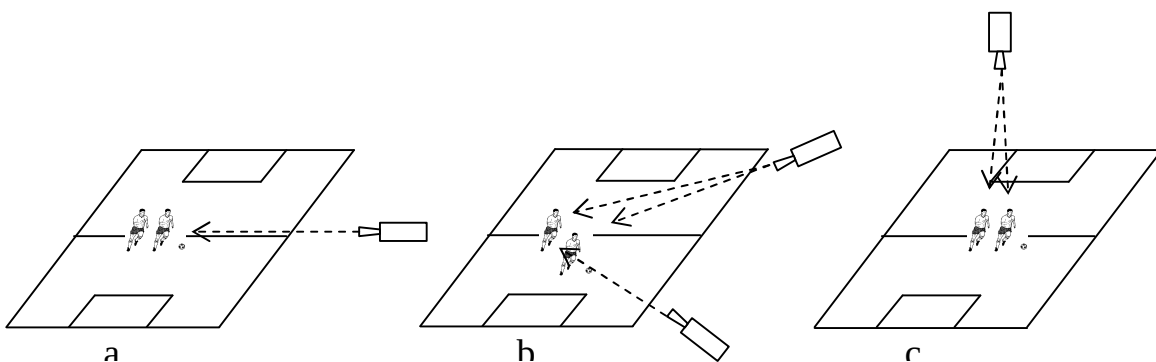


Abbildung 5. (a) zwei sich überlappende Spieler aus der Sicht einer Kamera in der Tribüne. Aus dieser Perspektive ist nur das vorstehende Objekt zu erkennen. (b) dieselbe Situation, aber gesehen von zwei verschiedenen Kameras in der Tribüne. Aus der Kombination der beiden Bilder ist die Ermittlung der beiden Objekte möglich, ist aber allerdings nicht trivial. (c) Wieder dieselbe Situation, aber von einer einzigen Kamera über der Spielfeldmitte aufgenommen. Beide Objekte sind gut erkennbar; die sichtbaren Flächen von den Spielern allerdings sehr klein.

Ein anderes Problem stellt aber die Positionierung der Kamera selbst – die Fussballspielfelder befinden sich normalerweise im Freien und ein Aufhängen der Kamera über das Spielfeld ist daher ohne spezielle Ausrüstung nicht

möglich. Diese Methode des Aufhängens ist bei anderen automatischen Erkennungssystemen sehr beliebt (wie z.B. das Vision-System von RoboCup), ist aber in diesem Fall nicht ohne weiteres machbar.

Ausserdem muss man noch ein anderes Problem beachten: die Objekte, gesehen von oben, werden noch kleiner. Die abgebildeten Flächen im Bild sind dann zu klein (auch im Vergleich zu nur einer Kamera von der Tribüne) und die Verfolgung wird schwieriger und ungenauer. Man kann ausserdem die Spieler nach ihren Nummern nicht mehr erkennen, da diese im Bild nicht zu sehen sind.

3. Verzerrung des Kamerabildes

Ein anderes Problem stellt die Perspektive und die Verzerrung der Kamera dar. Das Bild wird von der Optik der Videokamera stark verzerrt und die Berechnung der Koordinaten kann nicht mehr direkt erfolgen. Es muss also eine Nachbearbeitung der Daten stattfinden, die diese Transformation rückgängig macht.

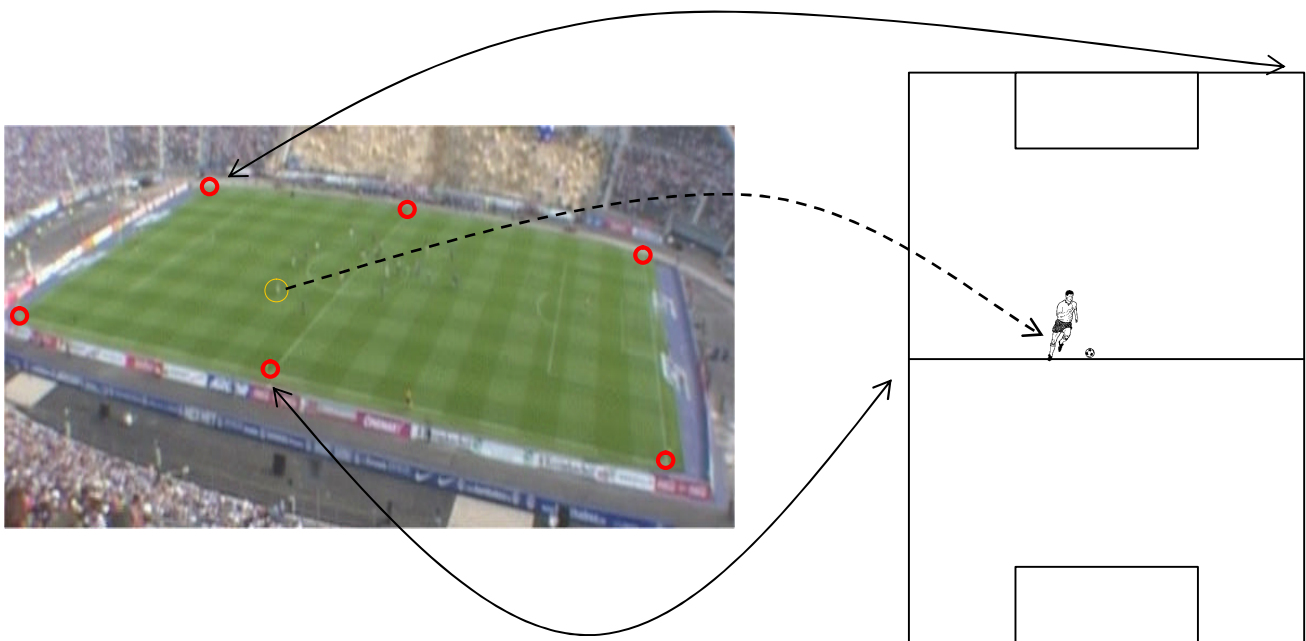


Abbildung 6. Ein Bild vom Spielfeld, aufgenommen von einer Kamera aus der Tribüne und die tatsächlichen Spielfeldkoordinaten nach der Entzerrung des Bildes.

Das Originalbild enthält das Spielfeld in einer bestimmten Perspektive, wie man auch von **Abbildung 6** sieht. Das Bild enthält ausserdem nicht nur das Spielfeld, sondern auch Teile von der Tribüne. Man muss also erstens das Spielfeld in die richtige Perspektive führen (wie im rechten Teil der **Abbildung 6** zu sehen ist), so dass es von oben gesehen wird und zweitens, muss man die radiale Verzerrung durch die Optik der Kamera dabei korrigieren. Diese radiale Verzerrung ist sehr gut im linken Teil der **Abbildung 6** zu sehen – die Punkte des oberen Randes des Spielfeldes befinden sich nicht auf einer Geraden, obwohl dies in der Realität der Fall ist.

Für die Transformation des Bildes (auch Kalibrierung genannt) gibt es verschiedene Methoden. Die einfachste dabei ist Referenzpunkte im Bild einzugeben (wie die roten Eckpunkten des Spielfeldes in der Abbildung). Diese Referenzpunkte haben im transformierten Bild festgelegte Koordinaten und die anderen Punkte im Bild können durch lineare Interpolation der Referenzpunkte berechnet werden. Für höhere Genauigkeit der Ergebnisse kann man statt lineare, auch quadratische Interpolation benutzen (die Referenzpunkte befinden sich auf einer Funktion zweiten Grades). Das Verfahren kann man anschaulich mit einem „Strecken“ und „Drehen“ der Referenzpunkte des Spielfeldes vergleichen.

4. Verfolgen der Spieler

Der entscheidende Teil des Systems ist die Verfolgung der Spieler und des Balls. Es geschieht auf der Basis von zwei wichtigen Informationen: die Farben des Spielfeldes im Gegensatz zu der Farbe der Trikots der Spieler und die relative Bewegung der Objekte in der Video-Sequenz.

Die Testbilddaten wurden von einer PAL - Kamera im sogenannten RGB-Raum aufgenommen, mit 720 x 576 Pixel pro Bild mit 25 Bilder (Frames) pro Sekunde. Der RGB-Raum stellt jede Farbe als ein Trippel mit ihren Anteilen von Rot, Grün und Blau dar.

Das Problem mit diesem Farbraum in dieser Anwendung ist aber die Intensität und die Sättigung einer Farbe. Da es sich bei den Daten um sehr viele Bilder handelt, die über längere Zeitabschnitte aufgenommen wurden (ein Fussballspiel dauert im Durchschnitt 90 Minuten), kann man nicht davon ausgehen, dass sich die Lichtverhältnisse im Stadion nicht ändern und dass die Helligkeit der Bilder dadurch nicht beeinflusst werden. Tatsächlich verändern sich diese Werte sogar von einem Bild zum anderen beträchtlich. Aus diesem Grund kann man vorher bestimmte Farben nicht benutzen, um die Spieler zu lokalisieren.

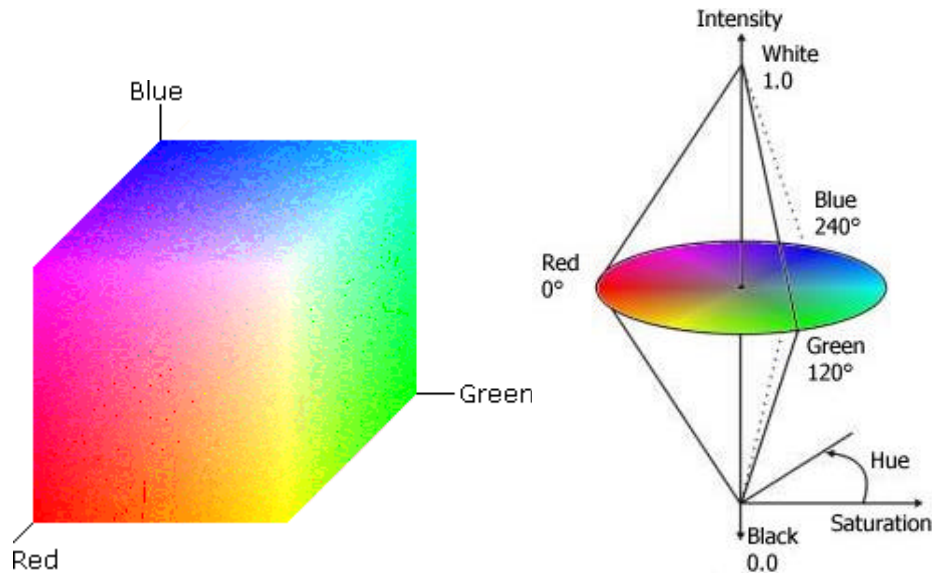


Abbildung 7. RGB-Raum (links) und HSI Raum (rechts)

Die Lösung findet man in einer Transformation des Farbenraumes RGB in den Farbenraum HSI. Man kann jede Farbe von RGB in HSI transformieren und rücktransformieren. Der HSI-Raum (Hue-Saturation-Intensity) stellt jede Farbe wieder als ein Trippel dar, diesmal aber als eine einheitliche Farbe, ihre Sättigung und Intensität. Daher kann man nur die reine Farbe benutzen und die Intensität und Sättigung, die so sehr von der Umwelt, Schatten usw. abhängen, nicht in Betracht ziehen.

Allgemein wird in RGB eine Farbe als ein Punkt im dreidimensionalen Raum dargestellt (wie in der **Abbildung 7** links zu sehen ist). Die drei Koordinaten sind genau die Anteile von Rot, Grün und Blau. In HSI wird jede Farbe dagegen als das Hue (Farbwinkel, wie man auch von der **Abbildung 7** rechts sehen kann), Sättigung (der Abstand eines Farbpunktes vom Zentrum des Raums) und Intensität dargestellt. Die Intensität ist in der Abbildung als eine dritte Achse zu sehen – man kann sich das Modell als einen dreidimensionalen Zylinder vorstellen, wobei die dritte Achse die Helligkeit (Intensität) ist und von 0 % (Schwarz) bis 100 % (Weiss) geht.

Das HSI Modell ist für die Verfolgung von farbig angezogenen Spielern sehr gut geeignet, da es, wie oben vorgestellt, die Farbe getrennt von den anderen Eigenschaften behandelt. Falls man aber auf Schwarz oder Weiss stösst, muss man zum RGB-Farbenraum zurückgreifen. Schwarz und Weiss sind nämlich keine Farben und ihr Hue ist daher im HSI-Modell nicht definiert.

In der Verfolgung der Spieler wird die Farbeigenschaft bei deren Suche auf dem Spielfeld benutzt. Man findet die Farbe einer „reinen Probe“ des Spielfelds - ein Bild mit dem leeren Spielfeld - und subtrahiert dann diese Farbe von allen Bildern, auf denen die Spieler gesucht werden (**Abbildung 8**). Damit bleiben auf dem untersuchten Bild nur die Pixel, die tatsächlich zu etwas anderem gehören als dem Spielfeld selbst (Spieler, Schiedsrichter, Ball).

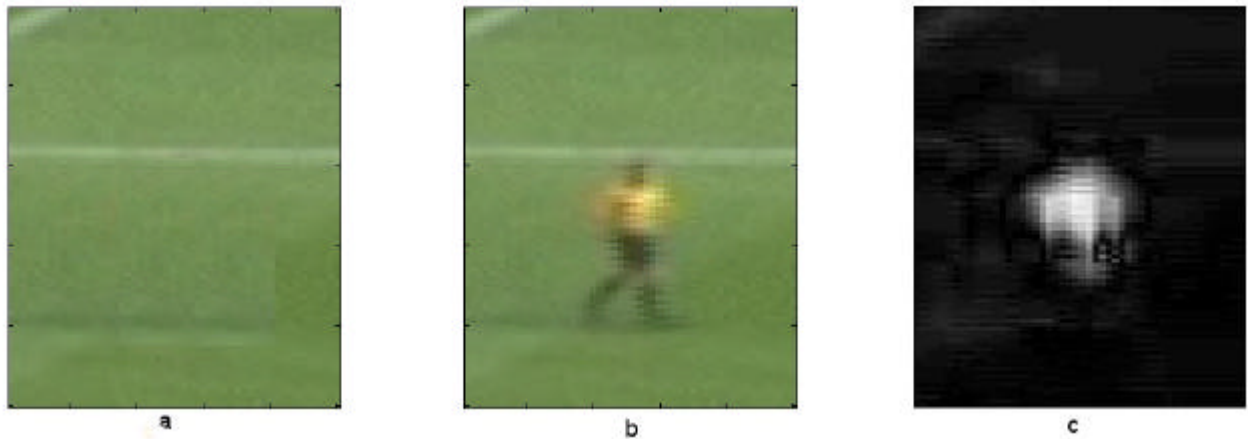


Abbildung 8. (a) Ein Bild mit dem leeren Spielfeld, aus dem die Farbe im HSI-Raum ermittelt wird. (b) Ein Bild mit einem Spieler auf dem Spielfeld (c) Das Ergebnis der Subtraktion der Hintergrund-Farbe aus dem zweiten Bild – es sind nur die Pixel geblieben, die zum Spieler gehören.

Zum eigentlichen Verfolgen der Objekte benutzt das System genau diese aus der Differenz mit dem Spielfeld-Hintergrund verbleibenden Pixel. Die Videoaufnahme erfolgt, wie oben erwähnt, mit 25 Bildern (Frames) pro Sekunde. Daraus folgt, dass sich ein Objekt (egal Spieler oder Ball) sich im nächsten Frame in unmittelbarer Nähe zu seiner alten Position befinden wird. Das System vergleicht die aus der Differenz mit dem Hintergrund stammenden Pixel mit denselben im nächsten Bild und berechnet den Abstand der beiden Objekte – auf diese Weise berechnet man auch direkt die Bewegung des Spielers.

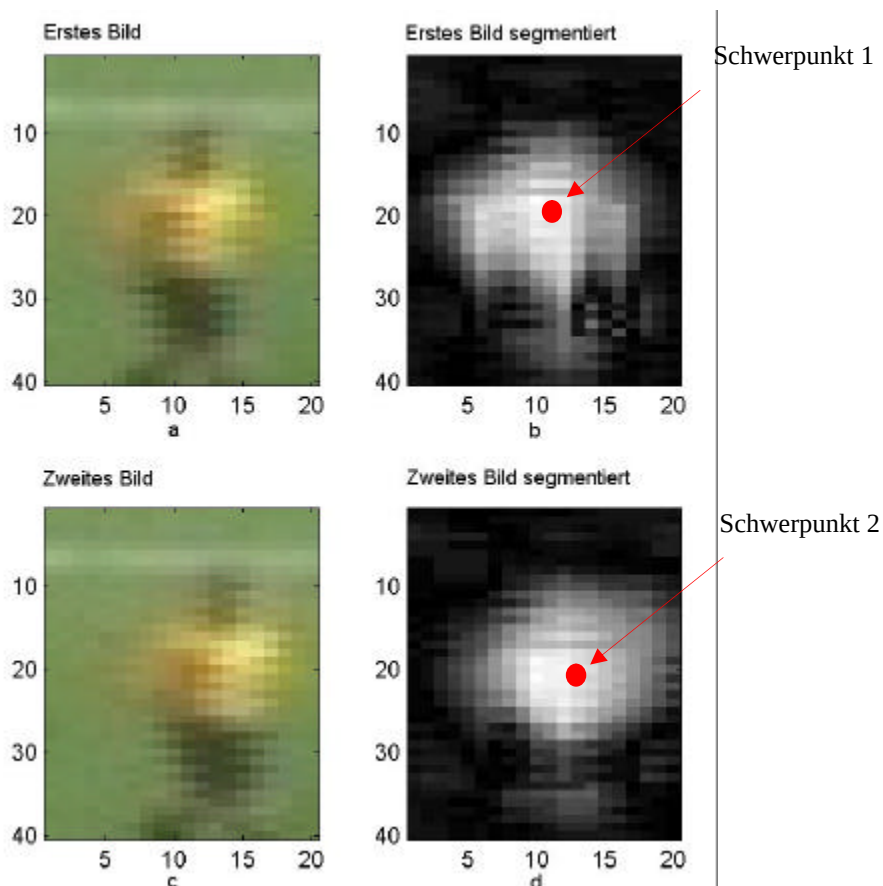


Abbildung 9. (a) Der Spieler im ersten Frame (b) die segmentierten Pixel vom ersten Frame und ihr Schwerpunkt (c) Derselbe Spieler ein Frame weiter und (d) die segmentierten Pixel mit ihrem Schwerpunkt vom zweiten Frame. Der Abstand der beiden Objekte und die berechnete Bewegung. Man berechnet zuerst den Schwerpunkt der beiden Objekte und findet dann den Abstand zwischen ihnen.

Man berechnet für das erste Frame der Bildsequenz den Schwerpunkt der segmentierten Pixel (die Pixel, die nach der Subtraktion geblieben sind) – den Mittelpunkt des Objekts als mittlere Position aller Pixel. Für das nachfolgende Frame segmentiert man wieder die Pixel, berechnet den Schwerpunkt und den Abstand zwischen den beiden Schwerpunkten im Bild (**Abbildung 9**). Genau dieser Unterschied ist die erfolgte Bewegung des Objekts (sein Schwerpunkt hat sich verschoben) und wird als Bewegung interpretiert. Das Verfahren benutzt den allgemeinen Schwerpunkt der segmentierten Pixel und ist daher resistent gegen Veränderungen in der Grösse der Objekte.

Die Initialisierung der einzelnen Objekte erfolgt manuell, durch einen Operator, der dem System in Zweifelsfällen am besten helfen kann. Er gibt dem System an, wieviele und welche Objekte er verfolgen möchte. Falls das System ein Objekt verliert (z.B. wegen Überdeckung) kann der Operator auch eingreifen und es wieder initialisieren.

Man sieht anhand von Abschnitten einer Sequenz in **Abbildung 10** wie das System einen Spieler verfolgt. Hier wurde das Verfahren für nur ein Objekt angewandt, für einen Zeitabschnitt von etwa 5 Sekunden, mitten im Spiel, wobei sich der zu verfolgende Spieler mit einem anderen überdeckt. Es wurde ausserdem nicht das ganze Spielfeld aufgenommen, sondern nur ein Teil davon, um die abgebildete Fläche der Spieler im Bild zu vergrössern.



Abbildung 10. Beispiel für die Verfolgung eines Spielers. (a) Initialisierung der Verfolgung (b) Ein anderer Spieler steht in unmittelbarer Nähe – trotzdem findet das System den richtigen Spieler (c) kurz nach der Trennung der beiden Spieler (d) derselbe Spieler 180 Frames (etwa 7 Sekunden) nach der Initialisierung.

Die Initialisierung erfolgt im ersten Frame (oben rechts im Bild) durch Anklicken des Objekts. Einige Frames weiter wird der Spieler immer noch verfolgt, überdeckt sich aber mit einem anderen Spieler. Im nächsten Frame sieht man,

wie die beiden wieder auseinandergehen und das System denselben Spieler weiter bis zum Ende der Testsequenz verfolgt.

Von diesen Testdaten sieht man, dass das System gut mit mehreren Problemen umgeht. Das Objekt ist klein, bewegt sich relativ schnell auf dem Spielfeld, ändert seine Richtung und wird von anderen Objekten überdeckt.

5. Anwendung der Ergebnisse

RoboCup an der Freien Universität Berlin

Das System wurde im Rahmen vom Forschungsprojekt „RoboCup – Fussballspielende Roboter“ am Institut der Informatik der Freien Universität Berlin entwickelt. In dessen Rahmen entwickeln wir eine Mannschaft von autonomen Roboter (die FU-Fighters), die gegen andere Mannschaften von ähnlichen Robotern Fussball spielen. Ein wichtiger Teil unserer Arbeit ist natürlich, den Robotern das nötige Verhalten beizubringen, d.h. für die Spiele werden Strategien und Taktik entworfen. Weil wir für die Roboter ein menschliches Vorbild haben, sind die besten Informationsquellen für die Entwicklung der Strategie echte Fussballspiele.

Das System „CatchIt!“ soll genau in diesem Gebiet eingreifen: Informationen aus echten Spielen für die Roboter-Entwickler sammeln, analysieren und Statistiken liefern. Diese Erkenntnisse werden uns dann weiterhelfen, die Strategie unserer Roboter-Mannschaft zu verbessern und schliesslich unsere Bemühungen gegen das Endziel von RoboCup zu bewegen: im Jahr 2050 gegen die Fussball-Weltmeister zu gewinnen (Abbildung 11)!



Abbildung 11. Bild von der letzten RoboCup-Weltmeisterschaft in Fukuoka, Japan. Humanoide Roboter spielen Fussball.

Interne Planung

Die Statistiken, die wir bei unseren FU-Fighters benutzen werden, sind auch für die Mannschaften selbst genauso interessant. Bei ihren Besprechungen und Planungen der Taktik und Strategie benutzen sie oft Videomaterial aus aufgenommenen Spielen. Sie analysieren sie, finden ihre Fehler, verbessern sich.

Mit Hilfe eines automatisierten Systems kann der Trainer die Analyse seiner Spiele und Spiele anderer Mannschaften einem Computer überlassen und sich voll auf die Interpretation der gewonnen Daten konzentrieren.

Unterhaltungs- und Informationsmedien

Alle Unterhaltungs- und Informationsmedien wie Fernsehen, Radios und verschiedene online-Anbieter versuchen heutzutage ihren Kunden so viele, so qualitative und interessante Informationen zu liefern, wie nur möglich. Sie bemühen sich ständig, ihre Angebote zu wechseln, auf die Wünsche ihrer Kunden anzugehen.

Bei den Sportspielübertragung (unabhängig von den benutzen Medien – Fernsehen, Radio oder Internet) bietet man schon lange nicht mehr nur eine trockene und uninteressante Darstellung des Spiels, sondern auch verschiedene Nebeninformationen. Im Fall von Fussballspielen sind das interne Informationen von den Trainern und den Mannschaften, Interviews, Statistiken usw.



Abbildung 12. Statistik erstellt für einen Spieler während des ganzen Spiels für die Anzahl der Ballkontakte.

Das System „CatchIt!“ bietet daher eine sehr kosteneffektive, einfache und effiziente Möglichkeit, Fussballspiele zu analysieren und Statistiken zu bilden. So kann man z.B. im Fernsehen eine vollständige Analyse von Spielen anbieten, die nur ein paar Stunden her sind. Und dafür braucht man eine oder zwei Kameras und einen normalen PC-Rechner.

Verschiedene mögliche Statistiken sind in den Abbildungen 12 bis 14 zu sehen, wie sie z.B. in den Sportnachrichten übertragen werden könnten.



Abbildung 13 . Statistik für die gelaufenen Kilometer von einem Spieler



Abbildung 14. Statistik für die mittlere Position eines Spielers

Komprimierung und Internet-Übertragung von Fussballspielen

Eine andere Anwendung des Systems stellt einerseits die Komprimierung und Speicherung der Ergebnisse der Analyse und andererseits die komprimierte Übertragung in Echtzeit eines Fussballspiels über langsamere Internet-Verbindungen dar.

Wie in der Einführung dieses Artikels vorgestellt, beträgt die Komprimierung eines ganzen Spiels etwa 1 : 60 000. Dies bedeutet eine enorme Speicherplatzoptimierung und viel leichtere Handhabung von den gewonnenen Informationen. Dabei enthält die komprimierte Version die ganze Information eines Spiels: die Positionen aller Objekte (Spieler, Ball, Schiedsrichter) für jede 1/25 Sekunde.

Man kann diese Komprimierung nicht nur für das Speichern der Spiele verwenden, sondern auch bei der Live-Übertragung. Viele Unterhaltungsmedien bieten eine Live-Übertragung per Internet, erreichen aber damit leider nur Kunden mit schnellem Internet-Zugang. Wenn man aber das System in Echtzeit nutzt und nur die Positionen der Spieler überträgt, spart man so viel Speicherplatz, dass auch die Internet-Nutzer mit den heutzutage langsamsten Verbindungen das Angebot nutzen können.

6. Künftige Forschung

Das System „CatchIt!“ verfügt im Moment über die Möglichkeiten, verschiedene Spieler und die Schiedsrichter zu verfolgen und ihre Trajektorien zu speichern. Grössere Probleme stellen aber die automatische Erkennung der einzelnen Spieler innerhalb einer Mannschaft und die Verfolgung des Balls dar.

Die Test-Aufnahmen wurden von einer Kamera ohne Stativ von der Tribüne aufgenommen. Das ganze aufgenommene Material hat eine Länge von 15 Minuten und eine Auflösung von 720 mal 576 Pixel pro Bild. Auf diese Weise kann man die unterschiedlichen Spieler verfolgen und das System als Ganzes testen, aber keine ausführliche Ergebnisse und Statistiken liefern.

Das erste Ziel meiner künftigen Arbeit ist die Verbesserung der Verfolgungsstrategie, damit auch der Ball lokalisiert werden kann. Die Probleme, die damit verbunden sind, ist einerseits die Grösse des Balls: in den Testdaten ist er nur noch wenige (2 bis 3) Pixel gross und kann vom Rauschen im Bild nicht unterschieden werden. Eine Lösung wäre eine Aufnahme mit grösserer Auflösung. Eine andere Lösung wäre aber auch, nach den Positionen der Spieler und ihre Bewegungsrichtungen die Position des Balls zu ermitteln.

Wie man auch in Abbildung 15 sieht, ist der Ball nur wenige Pixel gross. Dabei ist er im linken Teil des Bildes gut erkennbar und von den anderen Objekten separiert, aber im rechten Teil überdeckt er einen Spieler und kann vom System nicht mehr erkannt werden. Dieser Fall ist ein Zweifelsfall: ein Mensch kann den Ball immer noch finden, obwohl er mit den Farben des Spielers verschmolzen ist. Es existieren aber auch Beispiele, bei denen der Ball nicht mal von einem Menschen zu finden ist (wenn er z.B. hinter einem Spieler liegt).



Abbildung 15 . (a) ein Beispielbild, wo der Ball noch erkennbar gut zu sehen ist. (b) Ein Beispiel, wo der Ball überlappt mit einem Spieler ist und nicht mehr vom Spieler zu unterscheiden ist.

Falls zwei Spieler sich gegenseitig für eine gewisse Zeit im Bild überlappen, kann dies zum Vertauschen der beiden Spielern führen. In solchen Fällen muss dem System geholfen werden, die beiden Spieler wieder zu erkennen (eine Zweit-Initialisierung). Dies kann einerseits mit Hilfe von einem menschlichen Operator geschehen, aber auch andererseits mit einem automatischen Erkennungsverfahren, das die Nummern der Spielern ablesen kann.



Abbildung 16. Zwei Beispiele, bei denen man die Nummern der Spieler nicht identifizieren kann, obwohl sie von anderen Spielern nicht überlagert sind.

Dieses Erkennungsverfahren stellt wiederum einige Probleme dar: die Spieler stehen seitlich zur Kamera und ihre Nummern werden nicht abgebildet; die Nummern selbst sind zu sehen, aber durch Falten und Überlappungen nicht zu identifizieren (Abbildung 16). Die normalen Bilddaten, in denen das ganze Spielfeld aufgenommen wird, sind ausserdem nicht geeignet. Das System muss in Echtzeit funktionieren und über eine zweite Kamera verfügen, die auf die zu erkennenden Spieler zoomt und ihre Nummern sucht. Falls das Erkennungsverfahren über längere Zeit fehlschlägt, muss man indirekte Merkmale der Spieler in Betracht ziehen (Grösse, Frisur usw.), die zwar nicht eindeutig sind, aber in Zweifelsfällen oder in der Unterscheidung zwischen nur wenigen Spielern helfen können.

In Abbildung 16 werden zwei Beispiele gezeigt, wo das Erkennungsverfahren für die einzelnen Spieler fehlschlägt. Im linken Bild steht der Spieler mit dem Gesicht zur Kamera und verdeckt auf diese Weise die Sicht zu seiner Nummer auf dem Rücken. Im rechten Bild sehen wir einen Spieler, bei dem zwar die Nummer zu sehen ist, aber schwer zu identifizieren ist: erstens, weil sie durch Falten und Perspektive verzerrt wurde und mit einer 6 vertauscht werden könnte und zweitens, weil man nicht alle Ränder der Nummer gut erkennen kann und daher nicht weiss, ob es sich um eine Fünf oder eine Fünfzehn handelt.

Nicht an letzter Stelle ist zu erwähnen, dass auch neue und ausführliche Bilddaten erforderlich sind, um das System zu testen. Sie beinhalten am besten zwei Aufnahmen – einmal mit einer Kamera und einmal mit zwei Kameras, sind statisch (die Kamera bewegt sich nicht) und umfassen das ganze Spiel. Eine höhere Auflösung wäre auch wünschenswert, aber nicht unbedingt notwendig.

Ähnliche Arbeiten

In meinen Studien über ein automatisches Verfolgungssystem für Fussballspielen habe ich einige ähnliche Projekte entdeckt und teilweise benutzt (siehe Literaturquellen).

Ein sehr interessantes und erfolgreiches Beispiel ist die Arbeit von Paulo Cezar Pinto Carvalho, Flávio Szenberg und Marcelo Gattass von IMPA, Rio de Janeiro, Brasilien – Virtual Referee. Sie verfolgen dasselbe Ziel: anhand von Videosequenzen Fussballspiele zu analysieren.

Ein anderes Beispiel ist ein Projekt an der MTI Media Laboratory in Cambridge, USA – Computers Watching Football. Hier ist die Problematik ein bisschen anders: es wird amerikanischer Fussball gemeint, der andere Regel, Beschriftung des Spielfeldes usw. enthält.

Literaturverzeichnis

1. "An omnidirectional Vision System that finds and tracks color edges and blobs" , Felix v.Hundelshausen, Sven Behnke, Raul Rojas; Freie Universität Berlin, Institut für Informatik
2. „FU Fighters 2001 (Global Vision)“, Raul Rojas, Sven Behnke, Achim Liers, Lars Knippig; Freie Universität Berlin, Institut für Informatik
3. „Robust Real Time Color Tracking“, Mark Simon, Sven Behnke, Raul Rojas; Freie Universität Berlin, Insitut für Informatik
4. „FU Fighters Omni 2001 (Local Vision)“, Raul Rojas, Felix v. Hundelshausen, Sven Behnke, Bernhard Frötschl; Freie Universität Berlin, Insitut für Informatik
5. „Tracking People in Sport: Making Use of Partially Controlled Environment“, Janez Pers and Stanislav Kovacic; Faculty of Eletrical Engineering, University of Ljubljana, Slovenija
6. "Closed World Tracking", Stephen Intille, Aaron Bobick, Perceptual Computing Group, MIT Media Lab, Cambridge, USA.
7. "Image Based Modelling using a two-step Camera Calibration Method", Paulo Cezar Pinto Carvalho, Flávio Szenberg und Marcelo Gattass von IMPA, Rio de Janeiro, Brazilien
8. "Automatic Camera Calibration for Image Sequences of a Football Match", Paulo Cezar Pinto Carvalho, Flávio Szenberg und Marcelo Gattass von IMPA, Rio de Janeiro, Brazilien.
9. "Tsai's camera calibration method revisited", Berthold K.P. Horn