

# *blue-c*

## *Räumlich-immersive Projektions- und 3D Video-Portale für Telepräsenz*

### *Teil 1: Konzept, Hardware und Videoverarbeitung*

Markus Groß · Stephan Würmlin  
Martin Naef · Edouard Lamboray  
Christian Spagno

**Der folgende Beitrag beschreibt das Design und die Architektur von *blue-c*, einem Polyprojekt der ETH Zürich, an dem über einen Zeitraum von 4 Jahren bis zu 20 wissenschaftliche Mitarbeiter mitgewirkt haben und insgesamt 9 Doktorarbeiten entstanden sind.**

*blue-c* ist ein neuartiges, immersives Telepräsenzsystem, welches natürliches, kollaboratives Arbeiten räumlich getrennter Benutzer ermöglicht. Insbesondere erlaubt die Technologie, Kommunikationspartner realistisch und dreidimensional in eine

virtuelle Umgebung zu integrieren und mit dieser zu interagieren. *blue-c* kombiniert zum ersten Mal die simultane Verarbeitung multipler Videoströme mit fortschrittlicher Projektionstechnologie, so wie sie in bestehenden CAVE<sup>TM</sup>-Systemen eingesetzt wird, und erzeugt den Eindruck der vollständigen Immersion. Um *blue-c* zu verwirklichen, mussten zahlreiche Systemkomponenten hard- und softwareseitig von Grund auf entworfen und integriert werden. Besonders hohe Anforderungen ergaben sich hinsichtlich Synchronisation und Echtzeitverarbeitung der Daten. Die Architektur von *blue-c* ist skalierbar und erlaubt auch die Kombination mit weniger leistungsstarker Hardware.

tially Immersive Displays oder SIDs) signifikant zugenommen. Mittels multipler Projektoren erlauben solche SIDs dem Benutzer, in eine virtuelle Welt einzutauchen und mit ihr zu interagieren. Aus der Vielzahl von SID-Konzepten ist die so genannte CAVE<sup>TM</sup> [2] sicherlich mit das Bedeutendste. Hierbei wird mittels mehrerer Projektoren eine Rundumprojektion auf einem oftmals kubusförmigen Projektionszelt erzeugt. Stereoskopische Projektionen verstärken dabei die empfundene Immersion. Erweiterungen dieser Technologie erlauben es auch, Bilder multipler Projektoren mit gewöhnlichen Arbeitsumgebungen zu überlagern und somit die Gesamt-Projektionsfläche zu erhöhen. Beispiele sind PowerWalls oder die WorkWall<sup>TM</sup>. Ein weiterer Vertreter ist Elumens VisionDome, ein sphärisches Einzelprojektorsystem mit allerdings begrenzter Auflösung. Andere immersive Systeme sind z. B. die Responsive Workbench und die Holobench<sup>TM</sup>. Einige dieser Technologien sind bereits als Produkte erhältlich und werden auch in einem breiten Anwendungsspektrum, von Architektur über Kunst bis hin zur Medizin, eingesetzt.

Viele Anwendungen erfordern jedoch die Möglichkeit zur Telekollaboration, die in konventionellen CAVE-Systemen üblicherweise nicht gegeben ist.

#### Einleitung

#### Hintergrund und Motivation

In den letzten 10 Jahren hat die Bedeutung von räumlich-immersiven Projektionssystemen (Spa-

DOI 10.1007/s00287-005-0056-6  
© Springer-Verlag 2005

ETH Zürich  
Institute of Computational Science  
Haldeneggsteig 4  
8092 Zürich  
E-mail: grossm@inf.ethz.ch



**Abb. 1 Die zwei blue-c Portale in Aktion: Links ist das dreiseitige CAVE-System mit drei aktiven Projektionswänden und aktiver LED-Beleuchtung abgebildet. Das Bild rechts zeigt das einfachere PowerWall-System mit einer herkömmlichen Projektionswand. Beide Systeme akquirieren ein dreidimensionales Video der Benutzer mit insgesamt 16 Kameras und ermöglichen Stereo-Projektionen**

Daher ist man in den letzten Jahren bestrebt, den Aspekt der Telepräsenz mit SID-Technologie zu kombinieren. Oftmals werden dabei zweidimensionale Videokonferenzsysteme eingesetzt, welche ein Videobild des Kommunikationspartners in die virtuelle Szene einspielen. Solche Ansätze vermitteln allerdings nur einen sehr begrenzten Eindruck menschlicher Präsenz.

Somit sind Bestrebungen, eine volle dreidimensionale Integration des Benutzers zu erreichen, verständlich. Parallel zu diesen Aktivitäten intensivierte sich in den letzten Jahren auch die Forschung an dreidimensionalem Video oder 3D Video. Hierbei generieren Cluster von Kameras simultane Videostreams von unterschiedlichen räumlichen Positionen und erlauben damit eine freie Navigation in Raum und Zeit. Wir unterscheiden zwischen Ansätzen, die auf langwieriger Vorverarbeitung basieren wie z. B. Würmlin et al. [2002] und interaktiven Echtzeitsystemen wie Li et al. [20] oder Matusik et al. [8]. Das zentrale Problem ist hierbei die zuverlässige Extraktion von geometrischen Informationen.

Die Kombination von immersiver Projektion und Echtzeit-3D-Videoübertragung stellt eine sehr grosse technische Herausforderung in der systemorientierten Grundlagenforschung dar. Eines der umfassendsten Forschungsprogramme in den USA zu diesem Thema ist die *National Tele-Immersion Initiative* [16]. Ein weiteres Vorläuferprojekt zu diesem Thema ist auch das *Office of the Future* [14], wo man mit unsichtbarem strukturierten Licht eine Tiefenextraktion vornimmt. Eine vollständige Integration von immersiver Projektion und 3D-

Videoverarbeitung kann jedoch in keinem der erwähnten Systeme erreicht werden.

blue-c bietet eine Lösungsvariante für dieses schwierige Problem. Ziel war es, innovative Hard- und Softwaresysteme sowie Algorithmen zu entwickeln, welche eine Kombination von immersiver Projektion und 3D Video zum Zweck der Telepräsenz erlauben. Das blue-c Projekt war interdisziplinär mit Beiträgen aus der Informatik, der Informationstechnologie, dem Hardware-Design, der Konstruktionstechnik sowie der Architektur.

Die erreichten Forschungsergebnisse des Projektes können wie folgt zusammengefasst werden.

- Es wurde ein neuartiges Konzept für Projektionswände entwickelt, welches aktive Flüssigkristallwände verwendet [17]. Diese schalten schnell zwischen einem transparenten und opaken Zustand hin und her. Synchronisiert man derartige Displays mit Videokameras, können diese „durch die Wände hindurch“ sehen.
- Zur effizienten Verarbeitung der simultanen Videostreams musste ein 3D Video-Verarbeitungssystem entwickelt werden [22], welches den besonderen Bedürfnissen der blue-c Architektur Rechnung trägt. Sie verwendet 3D Punktwolken als fundamentales Videoprimitiv. Dies ermöglicht eine sehr effiziente Codierung, Datenübertragung und Darstellung.
- Aufgrund der sehr hohen Anforderungen an die Datenkommunikation sowie der unterschiedlichen, einem solchen System zu Grunde liegenden Datentypen musste eine entsprechende Netzwerk- und Kommunikationsarchitektur [5] entworfen werden.
- Schliesslich bedurfte es einer speziellen Anwendungsbenutzerschnittstelle (API) [12], welche

Applikationsprogrammierern komfortablen Zugang zu allen blue-c Funktionalitäten verschafft.

Im Folgenden geben wir zunächst einen Überblick über die Gesamtarchitektur von blue-c und beschreiben dann den Entwurf der Einzelkomponenten, inklusive der entwickelten Hardware (blue-c Hardwarekomponenten) sowie der Software und Methoden (3D Video-Verarbeitung, Kommunikation und Netzwerk, blue-c API). Weitere technische Details sind in den angegebenen Fachpublikationen zu finden oder auf der Projekt-Webseite: [blue-c.ethz.ch](http://blue-c.ethz.ch). Eine ausführliche Gesamtbeschreibung des Systems ist auch in [3] gegeben.

### Gesamtkonzept

Abbildung 2 zeigt die konzeptionellen Komponenten von blue-c. Insgesamt drei Doppel-LCD-Projektoren mit zusätzlichen Flüssigkristall-Blenden dienen zur aktiven Stereo-Bilderzeugung. Eine der hardwaretechnischen Kernkomponenten stellen die aktiv geschalteten Projektionswände dar, welche es den Kameras erlauben „durch die Wand zu sehen“. Wie wir noch beschreiben werden, löst dieses Konzept eine Reihe von technischen Problemen.

Zur kalibrierten Beleuchtung dienen ferner insgesamt 10 000 LEDs, welche synchronisiert mit den Kameras Licht in das System einspeisen. Um das Auge des Betrachters vor diesen Lichtpulsen zu schützen, wurden die für Stereoprojektion typischen LC-Stereobrillen in ihrer Arbeitsweise verändert. Zur Verarbeitung der insgesamt 16 Videoströme dient ein Linux PC-

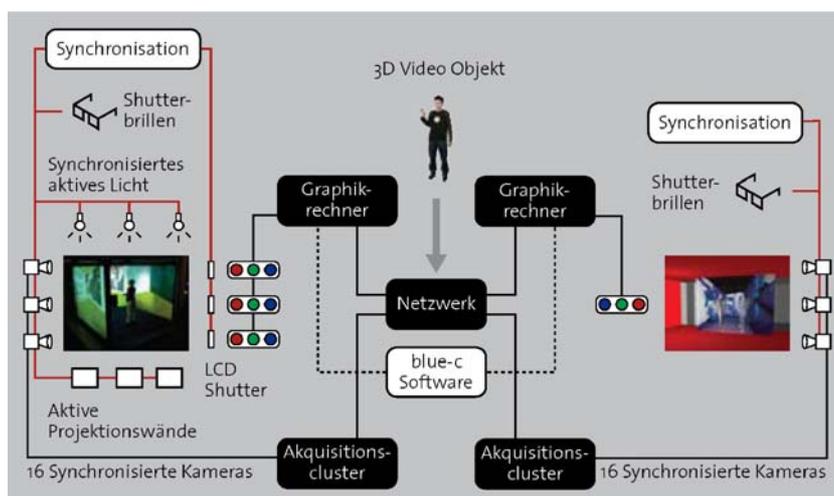
Cluster, die Bildgenerierung erfolgt mittels einer SGI Onyx 3200. Die präzise Synchronisation der beteiligten Hardware-Komponenten übernimmt eine eigens entwickelte Synchronisationshardware. Zusätzlich beherbergt die blue-c auch Hardware für Tracking, räumliche Audiosynthese, Sprachkommunikation und 3D-Benutzerinteraktion. Die Übertragung der 3D Videodaten verwendet ein neuartiges Format, die so genannten *3D Video-Fragmente*, welche sowohl auf effiziente Codierung als auch auf schnelles Rendering optimiert sind. Die beiden an der ETH gebauten Systeme sind über ein Gigabit-Ethernet, also ein IP-Netzwerk miteinander verbunden. Die auf Echtzeit-CORBA basierte Kommunikationsarchitektur verarbeitet eine Vielzahl hybrider Daten inklusive 3DVideo, Sprache, Events, Geometriedaten, Szenenaktualisierungen usw.

### blue-c Hardwarekomponenten

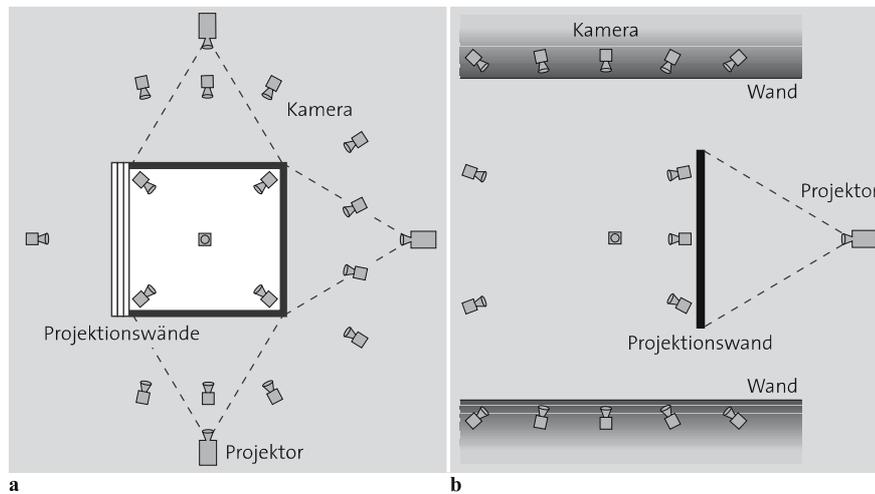
Das folgende Kapitel gibt einen Überblick über die Hardwarekomponenten von blue-c. Technisch detailliertere Beschreibungen sind in diversen Publikationen zu finden, insbesondere in Spagno und Kunz [17] sowie in der Dissertation von Spagno [15].

### Kameraanordnung

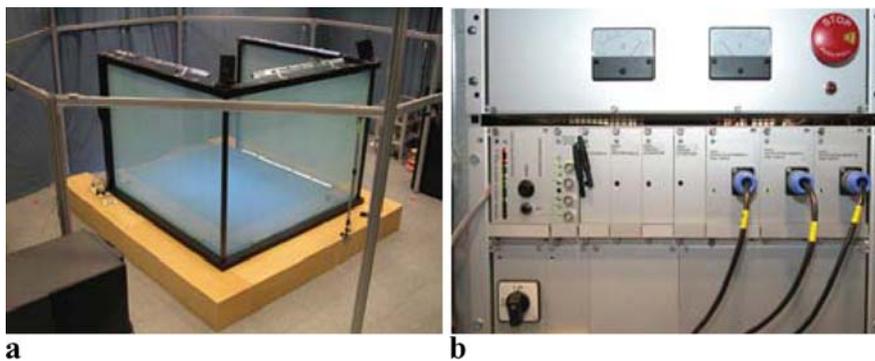
Das blue-c Hauptportal ist als dreiseitiges Stereorückprojektionssystem aufgebaut. Die planaren Projektionswände erleichtern die Verwendung multipler Projektoren. Ein zentrales Merkmal unseres Designs ist, dass sich die meisten Kameras ausserhalb des Projektionsraums befinden und gegenüber dem Benutzer von den Projektionswänden verdeckt



**Abb. 2 Gesamtkonzept von blue-c mit allen relevanten Hard- und Software-Komponenten. Die in rot dargestellten Linien repräsentieren die Synchronisation der einzelnen Hardwarekomponenten. Ebenfalls dargestellt sind die beiden Installationen ETH-Zentrum und ETH-Hönggerberg**



**Abb. 3 Illustration der Anordnung von Projektoren und Kameras in der blue-c. a) Hauptportal. b) Nebenportal**



**Abb. 4 a) reale Anordnung im Hauptportal. b) Synchronisationselektronik**

bleiben. Dies ist jedoch unumgänglich, denn eine Positionierung innerhalb des Aktionsraumes würde die empfundene Immersion erheblich stören und zu sehr grossen nichtlinearen Linsenverzerrungen führen. Das äussere Aluminium-Kameragerüst lässt hinreichend Freiräume für Positionsoptimierungen. Im Moment sind insgesamt 16 Dragonfly™ Firewire Kameras von Point Grey in VGA-Auflösung im Einsatz, 11 davon ausserhalb der Projektionswände. Die Positionen ergaben sich durch Softwaresimulation der Anordnung und haben wesentlichen Einfluss auf die Rekonstruktionsgüte. Schliesslich hüllt ein blauer Vorhang die gesamte Konstruktion ein.

Das Nebenportal am Standort ETH Höggerberg ist als interaktiver und multimedialer Raum im Departement Architektur frei zugänglich. Aufgebaut wurde eine passive Plexiglas-Projektionswand. 16 Firewire Kameras in VGA-Auflösung vom Typ 301fc der Firma Basler sind in die Wände und Decken des Raumes eingelassen und umrahmen die Projektionsfläche. Eine Übersicht der beiden Portale und ihrer Anordnung ist in Abb. 3 illustriert. Abb. 4 a)

zeigt die reale Anordnung des blue-c Hauptportals mit dem Kameragerüst.

### Aktive Projektionswände

Die aktiven Projektionswände stellen eine wesentliche Innovation des Hauptsystems dar. Zu diesem Zweck wurden umfangreiche Versuche mit Flüssigkristall-Glasscheiben verschiedener Hersteller durchgeführt. Derartige Scheiben sind eigentlich für innenarchitektonische Anwendungen entwickelt worden und für den statischen Betrieb ausgelegt. Nach ausgiebigem Test diverser Materialien haben wir schliesslich IsoLite™ Wände aus je 3 Segmenten aufgebaut. Durch Anlegen einer Spannung polarisiert sich der Flüssigkristall-Film und die Scheiben werden transparent. Allerdings stellen die Scheiben der Grösse 950 mm × 2240 mm und mit einem Gewicht von ca. 80 kg ein kapazitives Element von 11 µF dar, welches durch geeignete Ansteuer Elektronik kompensiert werden muss. Insbesondere die molekulare Relaxierung bei Spannungsabfall, also der Übergang von transparent nach opak, ist nicht extern beschleunigbar und setzt der oberen

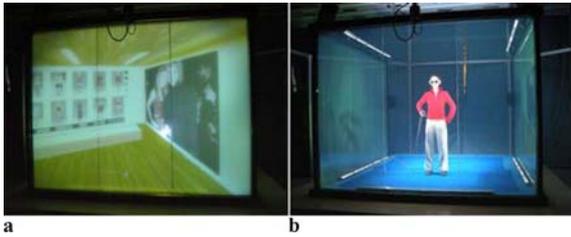


Abb. 5 Aktive Projektionswände a) opak geschaltet, b) transparent geschaltet

Taktfrequenz des Systems eine natürliche Grenze. Abbildung 5 zeigt die Projektionswände sowohl im opaken als auch im transparenten Zustand.

Zahlreiche Experimente [15] haben ergeben, dass eine Maximalfrequenz von 62,5 Hz (16 ms Periode) bei einem Taktverhältnis von 1 : 3,44 beste Resultate mit 70% Restopazität erzielt.

### Synchronisation und Zeitsteuerung

Die einzelnen Videobilder werden in einem kurzen Zeitfenster zwischen den Projektionssequenzen des linken und rechten Auges aufgenommen. Das prinzipielle Synchronisationsdiagramm dazu ist in Abb. 6 dargestellt. Die aktiven Projektionswände schalten in einem kurzen Intervall von 3,6 ms transparent und ermöglichen den synchronisierten Kameras eine ungehinderte Sicht auf den Benutzer. Während dieser Akquisitionsphase schliessen die aktiven Flüssigkristall-Blenden der Stereoprojektoren und verdecken damit das projizierte Bild, um Streueffekte zu verhindern. Ferner schliessen gleichzeitig auch die aktiven Flüssigkristall-Stereobrillen und schirmen das Licht der aktiven Beleuchtung vom Auge des Benutzers ab.

In der aktuellen Implementation von blue-c wird nur in jedem siebten Zyklus ein Videobild

aufgenommen, was zu einer Bildwiederholrate von insgesamt 8,9 Hz führt. Ebenso müssen die Verzögerungszeiten der beteiligten Hardware-Komponenten genau beachtet werden. Zur Synchronisation von Brille, Projektionswand, Projektoren und Lichtquellen dient ein innerhalb des Projektes entwickelter, programmierbarer Synchronisationsrechner auf Basis eines Mikrokontrollers [15]. Abbildung 4b) zeigt die Synchronisationshardware, welche in einem 19" Schrank eingebaut ist.

### Aktive Beleuchtung und Stereoprojektion

Ein grundsätzliches Problem bei der Kombination von Kameras und Projektoren stellen die gegensätzlichen Anforderungen an die Beleuchtungsbedingungen dar. Projektionen hoher Qualität bedürfen eines möglichst streulichtfreien Umfeldes, wogegen die Videoaufnahmen eine möglichst einheitliche, kalibrierte Beleuchtung benötigen. Diese verbessert die gesamte Videoverarbeitung, die 3D-Rekonstruktion sowie die Texturwiedergabe erheblich. Aus diesem Grund wurde eine externe Beleuchtung, bestehend aus mehreren Arrays mit insgesamt 10 000 einzelnen LEDs, entworfen. Die Lichtverhältnisse im Dunkel- und Hellzustand sind in Abb. 7 nochmals dargestellt. Die LED Reihen sind in Abb. 1 links unten dargestellt. Zur Bildgenerierung wurden LCD-Doppelprojektoren des Typs Sanyo XF12 LCD verwendet, welche über eine Helligkeit von je 3,500 ANSI-Lumen und über eine Auflösung von 1024 × 768 Pixeln verfügen. Die speziell angefertigten und vormontierten, ferroelektrischen Flüssigkristall-Blenden ermöglichen schliesslich

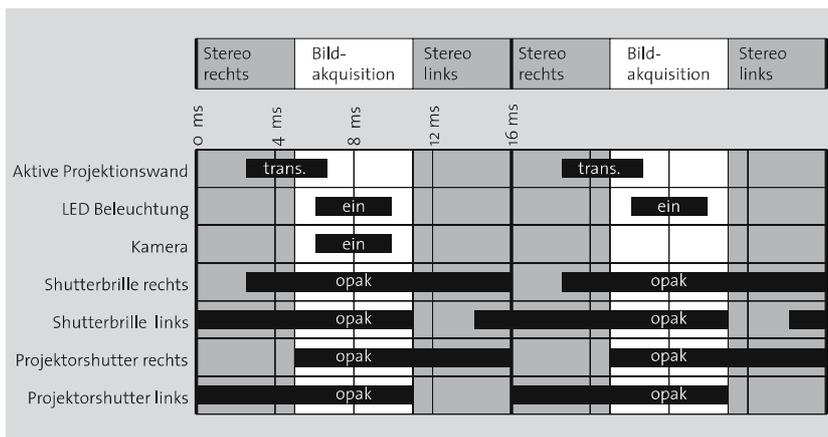
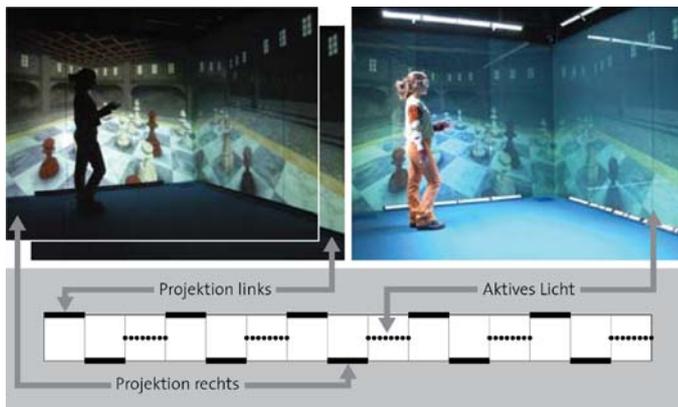


Abb. 6 Vereinfachtes Zeitdiagramm für die Synchronisation der beteiligten Hardware-Komponenten



**Abb. 7** Beleuchtungsbedingungen in der blue-c a) ohne und b) mit Zusatzbeleuchtung. Zu sehen ist auch die zeitliche Synchronisation der LED Lichtimpulse

einen qualitativ sehr hochwertigen Stereobetrieb mit sehr guter Kanaltrennung.

### 3D Video-Verarbeitung

Dieses Kapitel beschreibt die einzelnen Komponenten des 3DVideo-Systems der blue-c. Es verarbeitet die einzelnen 2D Videoströme in Echtzeit zu einer dreidimensionalen Repräsentation des Benutzers und überträgt diese über das unterliegende Netzwerk.

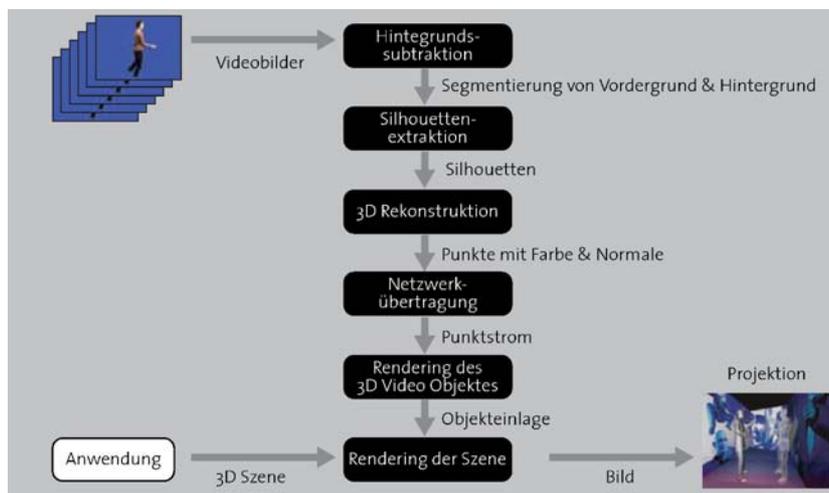
Eine umfassende technische Darstellung der 3D Videoverarbeitung ist auch in Würmlin et al. [21] sowie in der Dissertation von Würmlin [23] zu finden.

### Übersicht

Abbildung 8 zeigt die einzelnen Verarbeitungsschritte der Videopipeline der blue-c. In einem Vorverarbeitungsschritt muss das System zunächst sowohl farbvalenzmetrisch als auch geometrisch kalibriert werden. Zur Laufzeit werden die einzelnen Videoströme je einem Rechenknoten des

Akquisitionsclusters zugeführt. Dieser berechnet in einem ersten Schritt eine Trennung von Vordergrund und Hintergrund. Danach wird die Silhouette des Benutzers extrahiert. Die zweidimensionalen Silhouetten aus den 16 Einzelkameras werden einem Masterknoten zugeführt, welcher die eigentliche 3D Rekonstruktion durchführt. Aufgrund der hohen Anforderungen hinsichtlich Echtzeitverarbeitung verwendet man dazu ein Verfahren, welches die visuelle Hülle eines Objektes generiert [9].

Die zugrunde liegende Repräsentation der 3D Videodaten basiert auf *Punkt-Samples* [13]. Das dazu entwickelte Konzept der 3D Video-Fragmente erlaubt es, die räumliche und zeitliche Kohärenz des Videostroms zur Kompression und Übertragung zu nutzen. Schliesslich übernimmt ein geeigneter Punkt-Renderer [23] die Darstellung auf der Seite des Kommunikationspartners. Zusätzliche Funktionen umfassen einen 3D Video-Rekorder zur Aufnahme und Wiedergabe von Videosequenzen [20]. Diese Komponenten werden in den folgenden Abschnitten näher erläutert.



**Abb. 8** Einzelkomponenten und Verarbeitungsschritte der 3D Videopipeline von blue-c

## Systemkalibrierung

Die Extraktion von Geometrie aus Bildern ist eine der Grundaufgaben der Bildverarbeitung. Eine Grundvoraussetzung ist dabei die Kenntnis der sogenannten *extrinsischen* und *intrinsischen* Kameraparameter, welche die Abbildungsgeometrie und räumliche Anordnung der Kameras beschreiben. Da diese vorab nicht bekannt sind, müssen sie in einem Kalibrierungsschritt berechnet werden. Aufgrund der expliziten Synchronisation der einzelnen Kameras dient dazu ein recht einfaches Verfahren, bei welchem eine Person bei laufender Kamera mittels eines Laserpointers das Zielvolumen mit virtuellen Punkten ausfüllt. Aufgrund der geringen Distanz der Kameras vom Benutzer (2.5 m), ergeben sich auch radiale Linsenverzerrungen, welche ebenfalls auskorrigiert werden müssen. Eine detaillierte Beschreibung der geometrischen Kalibrierung von blue-c kann in [19] gefunden werden.

Obwohl homogene Kamerasysteme in blue-c im Einsatz sind, sind die Farbvalenzen nicht genügend homogen für unser 3D Videosystem. Ebenso führen leichte, positionsabhängige Unterschiede in der Restkapazität der Projektionswände zu Farbverzerrungen, welche ebenso auskorrigiert werden müssen. Dies erfolgt durch eine lineare Transformation der Eingangsfarben der einzelnen Kameras. Die Matrixeinträge errechnen sich aus Referenzkameras, welche im Voraus definiert werden. Im Falle des Hauptportals sind dies die Kameras im Innern der blue-c.

## Hintergrundtrennung und Silhouetten

**Hintergrundtrennung.** Trotz der Vielzahl der existierenden Segmentierungsverfahren zur Trennung von Vordergrund und Hintergrund, schränkt die Echtzeitanforderung die Wahlfreiheit erheblich ein. Der verwendete Algorithmus basiert auf einer einfachen statistischen Beschreibung des Hintergrundes, welche das System in einem initialen Trainingsschritt lernt. Der Vergleich der Pixelstatistik einer kleinen räumlichen Nachbarschaft mit der Hintergrundstatistik lässt eine Hypothese über die Zugehörigkeit des jeweiligen Pixels zu. Zusätzlich ist das Hauptportal in einen blauen Vorhang (siehe Abb. 4 a) eingehüllt, welcher den Hintergrund statistisch homogenisiert. Die Wände und Decken des Nebenportals sind in einem homogenen roten Farbton gehalten. In Abb. 9 ist ein typisches Video-

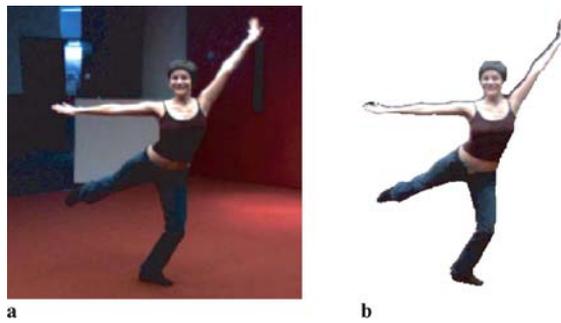


Abb. 9 a) Original-Videobild. b) Hintergrundtrennung

bild (a) vor und (b) nach der Hintergrundtrennung gezeigt.

**Silhouetten.** Die Qualität der 3D Rekonstruktion und damit des resultierenden 3D Video hängt neben der Hintergrundtrennung vor allem von der Anzahl der möglichen Geometriepunkte und damit von der Auflösung der Silhouette ab. Zum Zweck der adaptiven Auflösungssteuerung haben wir eine Multiskalenrepräsentation entwickelt, welche die Silhouette in sukzessiv höherer Auflösung, ausgehend vom segmentierten Bild, berechnet. Das Verfahren unterteilt das Bild in Blöcke zusammenhängender Pixel, wobei die Blockgröße durch einen Parameter  $\alpha$  vom Benutzer gesteuert werden kann. Im praktischen Betrieb setzen wir  $\alpha = 1$  oder  $2$  und erhalten  $2 \times 2$  oder  $4 \times 4$  Gitter.

## 3D Rekonstruktion und 3D Video Repräsentation

Konzeptionell betrachtet generalisiert die von uns entwickelte 3D Video-Repräsentation die üblichen zweidimensionalen Bildpixel hin zu Punkten im Raum, den so genannten *3D Video-Fragmenten*. Dieser Ansatz kombiniert die strukturelle Einfachheit und Regularität eines Bildpixels mit den Leistungsmerkmalen komplexerer, dreidimensionaler Repräsentationen. Ein Videofragment ist also ein Abtastwert auf der Oberfläche eines 3D Video-Objektes, welcher neben geometrischen Attributen, wie Position und Normale, auch Darstellungsattribute, wie Farbe und Textur, vorsieht. Eine Konnektivität zwischen einzelnen Fragmenten wird dabei nicht berücksichtigt. Dies vereinfacht die Kompression sowie die Übertragung erheblich. Abbildung 10 veranschaulicht die Beziehung zwischen den 2D Pixeln und den 3D Video-Fragmenten.

**Differentielle Übertragung.** Um die räumlich-zeitliche Kohärenz dieser Videofragmente auszunutzen, wurden die folgenden Fragment-Operatoren definiert:

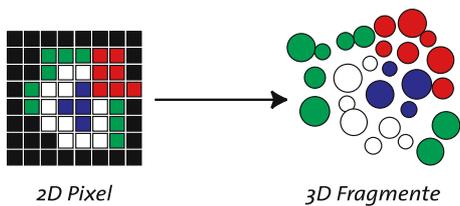


Abb. 10 Beziehung zwischen 2D Pixeln und 3D Video-Fragmenten

- **INSERT** fügt ein neues 3D-Fragment in die bestehende Repräsentation ein, nachdem es in einer Kamera sichtbar geworden ist.
- **DELETE** löscht ein unsichtbar gewordenes Element aus der Repräsentation.
- **UPDATE** korrigiert die Attribute für Geometrie oder Darstellung eines Fragments aufgrund aktueller Videoinformationen.

Die Sequenz dieser Operatoren erzeugen einen differenziellen Fragment-Operatoren-Strom, welcher eine 3D Datenstruktur beim Kommunikationspartner unterhält. Ein INSERT-Operator resultiert aus der Rückprojektion eines Pixels mit Farbattributen vom Bildraum in den dreidimensionalen Objektraum. Dazu wird ein 3D Rekonstruktionsalgorithmus, basierend auf der visuellen Hülle, angewendet [9], welcher in Echtzeit Tiefen- und Normaleninformation berechnet. Die Fragmentprimitive weisen eine bijektive Abbildung von Tiefen- und Farbwerten auf.

UPDATE-Operatoren errechnen sich aufgrund aller Pixel, welche in vorherigen Videobildern eingefügt wurden und sich immer noch im Vordergrund befinden. UPDATE-Operatoren werden ferner in drei Kategorien unterteilt: Die Erkennung von Farbveränderungen wird zwischen zwei Videobildern ausgeführt und hat einen UPDATECOL-Operator zur Folge. UPDATEPOS-Operatoren sorgen für Veränderungen an der Geometrie und werden auf räumlich zusammenhängenden Bereichen von Bildpixeln analysiert. Für bestimmte Pixel eines

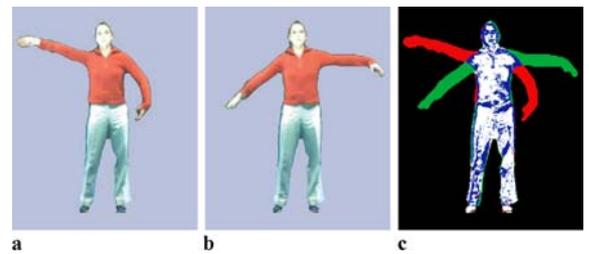


Abb. 11 Die Berechnungen der 2D Pixel-Operatoren: a) Silhouette zur Zeit  $t_{i-1}$ . b) Silhouette zur Zeit  $t_1$ . c) Klassifikation der Bildpixel gemäss Beschreibung im Text. Rot steht für expired, grün für new, blau für color changed, weiss für color unchanged und schwarz für Hintergrund

Bereiches werden neue Tiefenwerte berechnet. Nur wenn die Differenz zu bisherigen Tiefenwerten einen Grenzwert überschreitet, wird die 3D Information für den ganzen Bereich neu berechnet. Der explizite Test aller Pixel ist aufgrund der Rechenkosten für die 3D Rekonstruktion nicht möglich. Unsere Methode stellt aber eine effizient Lösung für das Problem der unkorrelierten Textur- und Tiefenbewegungsfelder dar. Positions- und Farbveränderungen führen zum kombinierten UPDATEPOSCOL-Operator. Alle anderen Pixel, welche für UPDATES vorgemerkt wurden, bleiben unverändert und keine weitere Bearbeitung ist notwendig. Ferner verwendet blue-c ein Verfahren zur Bewegungsschätzung im Bildraum, welches Farb- und Positionsveränderungen erkennt und daraus 3D Fragment-Operatoren ableitet. Eine Klassifikation des Videobildes einer der Kameras ist in Abb. 11 gezeigt.

**Dynamische Kamerakontrolle und Texturdetails.** Eine erste Massnahme zur Beschleunigung der 3D Video-Berechnungen für die aktuelle Kameraposition des Kommunikationspartners ist die dynamische Auswahl einer Untermenge von aktiven Kameras, typischerweise basierend auf ihrer Distanz von dieser Position. Wir definieren für jede Kamera einen so genannten *Aktivitätslevel* und gewichten dabei die Beiträge der einzelnen Kameras, gemäss eines Interpolationsverfahrens

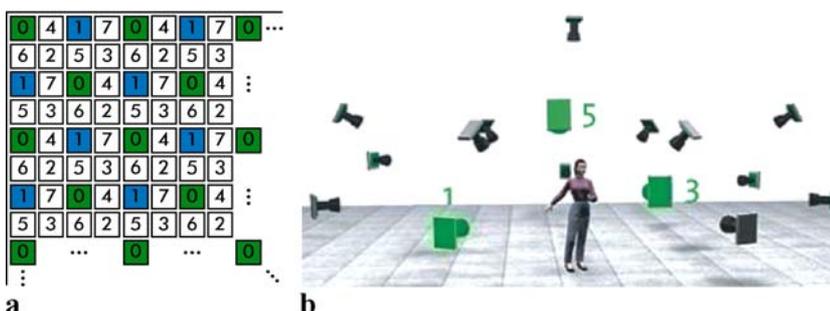


Abb. 12 Texturniveaus: a) Abtastmuster für verschiedene Aktivitätsstufen, b) Beispiel für 3 aktive Kameras auf unterschiedlichen Niveaus, gesehen von einer virtuellen Kamera aus

aus. Dabei unterscheidet man zwischen Kameras, die nur für die geometrische Rekonstruktion aktiv (*geometry active*) sind, sowie Kameras, welche zur Berechnung der Darstellungsattribute Informationen liefern (*texture active*). Üblicherweise ist die Anzahl der geometrieaktiven Kameras höher als diejenige der texturaktiven Kameras. Zur Implementation dieser Funktionalität muss das Kommunikationsgerüst (Abschnitt Kommunikation und Netzwerk) die virtuelle Kameraposition des Kommunikationspartners über das Netzwerk zurückliefern.

Eine zweite Massnahme zur dynamischen Adaptierung des Berechnungsaufwandes pro Knoten sind so genannte *Texturniveaus*. Diese definieren die Anzahl der 2D Bildelemente einer bestimmten Kamera, welche der 3D Rekonstruktion jeweils zugrunde gelegt sind. Das Texturniveau bestimmt damit die aktuelle Auflösung, mit der das Videobild verarbeitet wird. Die initialen Gewichte berechnen sich gemäss der oben angegebenen Distanzgewich-

tung. Zusätzlich verwenden wir zur Berechnung des Texturniveaus ein Mass für die aktuelle Systemauslastung des jeweiligen Rechner-Knotens für die 3D Rekonstruktion. Ist also der Berechnungsaufwand zu gross, so reduziert sich die Bildauflösung, was wiederum zu einer Reduktion des Berechnungsaufwandes führt. Dies implementiert eine einfache, lineare Regelung zur Stabilisierung der Bildwiederholraten. Das Bildabtastmuster für verschiedene Texturniveaus ist in Abb. 12 veranschaulicht.

Jede Kamera ist mit einem PC-Knoten verbunden, welcher die beschriebene Videopipeline berechnet. Ein zentraler Masterknoten übernimmt schliesslich die silhouettenbasierte 3D Rekonstruktion. Die für einen bestimmten Zeitschritt jeweils aktuelle Repräsentation wird vollständig auf der Empfängerseite gehalten und durch die beschriebenen Operatoren entsprechend aktualisiert. Die dazu entwickelte Datenstruktur erlaubt einen zeitgleichen, asynchronen Lese- und Schreibzugriff.