

# ENTWURF UND IMPLEMENTIERUNG EINER SYSTEMPLATTFORM FÜR ECHTZEIT-BILDVERARBEITUNGSANWENDUNGEN AN BORD EINES VTOL-MAVS

C. Kappenberger

DLR e.V. Institut für Flugsystemtechnik, 38108 Braunschweig, Deutschland

## Zusammenfassung

Der vorliegende Beitrag behandelt den Entwurf sowie die Realisierung einer DSP-basierten Plattform zur Unterstützung der autonomen Navigation eines *Vertical Take-Off and Landing Micro Aerial Vehicles* (VTOL-MAV) mittels Bildverarbeitungsalgorithmen. Zunächst werden die erforderlichen Entwurfsziele benannt sowie die einschränkenden Randbedingungen erläutert. Die eingesetzten Hardwarekomponenten für die Bilderfassung und -verarbeitung sowie die Datenübertragung und Speicherung werden vorgestellt sowie die entwickelte Softwareumgebung, unter besonderer Berücksichtigung der Anforderungen an eine Echtzeitverarbeitung der Bilddaten, beschrieben. Um die Leistungsfähigkeit der Systemplattform zu zeigen, werden zwei implementierte Anwendungen der Bildverarbeitung vorgestellt und die Unterschiede zu einer PC-gestützten Lösung aufgezeigt.

## 1. EINLEITUNG

Fortschritte in den Bereichen der Mikroelektronik, der elektrischen Antriebstechnik sowie insbesondere auch bei modernen Akkumulatortechnologien haben in den letzten Jahren die Entwicklung einer neuen Form von Luftfahrzeugen für die Luftbildfotografie und Fernerkundung möglich gemacht, die so genannten *Micro Aerial Vehicles*. Hierbei handelt es sich um kleine, unbemannte Luftfahrzeuge mit einem Gewicht von unter einem Kilogramm. MAVs zählen zur Gruppe der *Unmanned Aerial Vehicles* und können als Starr- oder Drehflügler ausgeführt sein, in jüngerer Zeit befinden sich vermehrt auch hybride Formen [1] in Entwicklung, die das Flugverhalten von Insekten imitieren.

Da insbesondere Drehflügel-MAVs ungeregelt ein instabiles Flugverhalten aufweisen, werden verschiedene unterstützende Module wie Beschleunigungs- und Drehratensensoren, Radar- und Laserabstandsmesser sowie barometrische Höhenmesser eingesetzt, um die Fluglage zu stabilisieren. Durch Erweiterung um satellitenbasierte Navigationssysteme wie das *Global Positioning System* ist es durch Fusion der verschiedenen Sensordaten möglich, eine präzise Positionsbestimmung des MAVs vorzunehmen und somit autonom zu navigieren. Neben diesen Sensoren bieten auch automatisierte Verfahren zur Bildverarbeitung die Möglichkeit, die Positionsbestimmung zu unterstützen. Darüber hinaus lassen sich diese Verfahren als eigenständige Anwendung zur Erweiterung des Einsatzspektrums von MAVs einsetzen, beispielsweise in der Objekterkennung oder zur Georeferenzierung in der Kartografie.

Aufgrund des hohen Datenvolumens, das bei der automatisierten Bildverarbeitung anfällt, ergeben sich Mindestanforderungen an die Systemumgebung hinsichtlich Speicherplatz und Verarbeitungsgeschwindigkeit. Mit zunehmender

Leistungsfähigkeit einer Hardwareplattform steigen jedoch auch Leistungsaufnahme, Masse und Abmessungen; dies reduziert die Möglichkeiten einer Integration in einem MAV.

## 2. RANDBEDINGUNGEN UND ENTWURFSZIELE

Das Ziel der Arbeit ist es, eine Plattform zur Integration von Bildverarbeitungsanwendungen an Bord des VTOL-MAVs *MD4-200* [2] zu realisieren. Bei diesem MAV (Bild 1) handelt es sich um einen Quadrokopter, der am *Institut für Theoretische Elektrotechnik und Systemoptimierung* in Zusammenarbeit mit der Firma *Microdrones GmbH* entwickelt wurde



BILD 1. VTOL-MAV MD4-200

Dieser Quadrokopter findet Anwendung in erster Linie in der Luftbildfotografie und wird auch zunehmend von Einsatz- und Rettungskräften verwendet, um am Einsatzort einen detaillierten Lageüberblick zu erhalten. Um den Benutzer weitestgehend von Aufgaben der Flugsteuerung und Navigation zu entbinden, verfügt der MD4-200 über eine Avionik, die über Beschleunigungs-, Drehraten-, Druck- und Magnetfeldsensoren sowie einen

GPS-Empfänger eine Stabilisierung der Fluglage sowie eine weitgehend autonome Navigation ermöglicht.

Der Antrieb des MAVs erfolgt elektrisch über 4 bürstenlose Gleichstrommotoren, die Stromversorgung über einen an Bord befindlichen Lithium-Ionen Akkumulator. Die Systemplattform zur Bildverarbeitung soll vollständig in diese Umgebung unter Berücksichtigung der gegebenen Randbedingungen (Tabelle 1) integriert werden. Erstrebenswert ist hierbei insbesondere ein möglichst geringes Gewicht, um die Nutzlast des Quadropters sowie die maximale Flugdauer, die abhängig von der Gesamtmasse ist, in nur geringem Maße einzuschränken.

Abmessungen	54,0 cm x 54,0 cm x 19,9 cm
Eigengewicht	800 g
Abmessungen Nutzlast	11,0 cm x 13,0 cm x 8,0 cm
Nutzlastgewicht	200 g
Flugdauer	20 Minuten
Aktionsradius	5 km

TAB 1. Technische Daten MD4-200

Neben der Erfüllung der durch die fliegende Plattform vorgegebenen Randbedingungen soll das System zur Bildverarbeitung die folgenden Funktionen ermöglichen:

- Aufnahme von Standbildern und Videosequenzen
- Bildauflösung mindestens 1024 Pixel x 768 Zeilen (1 Megapixel)
- Speicherung der Bilddaten zur späteren Verarbeitung im JPEG- und RAW-Format auf einem FAT-Dateisystem
- Echtzeitübertragung an Bodenstation über vorhandene Funkschnittstelle in TV-Norm
- Aufnahme und Verarbeitung von stereographischen Bilddaten
- Einsatz ausschließlich digitaler Datenpfade zur Erhaltung der Bildqualität
- Einfache Migration von vorhandenen Algorithmen auf die Echtzeit-Umgebung

Neben dieser Funktionalität sind Schnittstellen zu den bestehenden MAV-Systemen wie Navigations- und Flugrechner sowie zur Echtzeit-Interaktion mit dem Benutzer zu realisieren.

### 3. SYSTEMPLATTFORM

Die Realisierung der Systemplattform gliedert sich im Wesentlichen in die Bestandteile Hardwareplattform und Softwareumgebung.

#### 3.1. Hardware

Aufgrund der in Kapitel 2 skizzierten technischen Einschränkungen bezüglich Abmessungen, Masse und Leistungsaufnahme ist der Einsatz einer PC-basierten Umgebung nicht möglich und der Entwurf eines eingebetteten Systems erforderlich. Um die erforderliche Leistungsfähigkeit zur Realisierung von Anwendungen der Bildverarbeitung zu erreichen, wird bei der vorgestellten Hardwareplattform ein digitaler Signalprozessor (DSP) eingesetzt. DSPs sind insbesondere darauf ausgelegt, hohe Datenvolumina in Echtzeit verarbeiten zu können. Echtzeit bedeutet in diesem Kontext, dass die

Verarbeitung eines Datenblocks innerhalb einer definierten Zeit abgeschlossen ist.

#### 3.1.1. Digitaler Signalprozessor

Verwendung findet ein DSP des Typs ADSP-BF561 aus der Blackfin-Baureihe der Firma Analog Devices. Dieser verfügt neben zwei parallelen CPU-Kernen über eine Reihe von Funktionen, die den Einsatz in Applikationen der Bildverarbeitung unterstützen. Hierzu gehören insbesondere zwei *Parallel Peripheral Interfaces* (PPI), die die Integration von Hardware-Videokomponenten wie Kamerasensoren ohne komplexe Schnittstellenbausteine ermöglichen. Eine PPI-Schnittstelle bietet die Möglichkeit, einen bis zu 16 Bit breiten Datenbus synchron zu einer externen Taktquelle sowie zwei weiteren Synchronisationssignalen, wie sie häufig bei gängigen Videoschnittstellen zur Signalisierung des Beginns einer Zeile beziehungsweise eines Bildes eingesetzt werden, einzulesen oder auszugeben.

Zur Reduktion des Entwicklungsaufwands wird ein Prozessormodul der Firma *Bluetechnix* eingesetzt, das bereits die wesentlichen Komponenten (Tabelle 2) wie Arbeitsspeicher und Stromversorgung für den Betrieb des ADSP-BF561 enthält.

Abmessungen	3,6 cm x 3,1 cm x 0,5 cm
Masse	5 g
Flüchtiger Speicher	64 MB SDRAM
Nichtflüchtiger Speicher	8 MB NOR-Flash
Stromaufnahme	Maximal 470 mA bei 3,3 Volt

TAB 2. Technische Daten Prozessormodul Bluetechnix CM-BF561

#### 3.1.2. Bildsensoren

Als Bildsensoren finden zwei identische Kameramodule des Typs OV2630 [3] der Firma Omnivision Verwendung. Es handelt sich hierbei um konfektionierte Sensorchips, die neben einem Objektiv mit fester Brennweite über einen Flachbandkabelanschluss verfügen, der eine in gewissen Grenzen flexible Montage der Kameramodule an der Systemplattform ermöglicht.

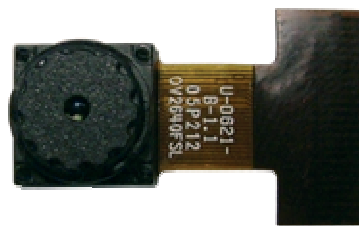


BILD 2. CMOS Kameramodul OV2630

Die Ausgabe der Bilddaten erfolgt kontinuierlich als eine Folge von 8- oder 10-Bit Helligkeitswerten über eine parallele LVTTTL Schnittstelle, die Farbkodierung erfolgt dabei nach dem Bayer-Schema. Zur Synchronisation werden durch die Kameramodule sowohl ein Pixeltakt als auch die zwei Signale HSYNC und VSYNC generiert, die den Beginn einer neuen Bildzeile beziehungsweise eines neuen Bildes markieren.

Neben der Schnittstelle zur Übertragung von Bilddaten verfügen die Kameras über jeweils einen Serial Camera Control Bus (SCCB), der weitestgehend nach dem *Inter-*

I<sup>2</sup>C (I2C) Protokoll arbeitet und über den die Betriebsparameter des Sensorchips konfiguriert werden können. Diese Parameter definieren Werte wie beispielsweise Bildauflösung, Wiederholrate oder Belichtungsdauer.

Der Anschluss der Kameramodule erfolgt an die erste PPI-Schnittstelle (PPI0) des ADSP-BF561. Da dieser nur über eine Wortbreite von 16 Bit verfügt, können die Bilddaten nur mit einer Breite von 8 Bit ausgelesen werden. Alternativ ist die Verwendung jeweils eines PPI-Anschlusses pro Kameramodul möglich, bei dieser Variante stünde allerdings kein Videoausgang zur Verfügung. Ein weiterer Vorteil des gewählten Verfahrens besteht darin, dass die Bilddaten beider Kameras pixelsynchron ausgelesen werden können. Die Synchronisation der Bildabtastung beider Kameras ermöglicht den Einsatz der Bildverarbeitungsplattform in der Stereobildverarbeitung. Erfolgt die Erfassung der beiden Einzelbilder nicht synchron, ergeben sich Differenzen zwischen beiden Perspektiven, die neben geometrischen auch zeitliche Ursachen haben und somit zusätzliche Fehlerquellen in den jeweiligen Algorithmen induzieren.

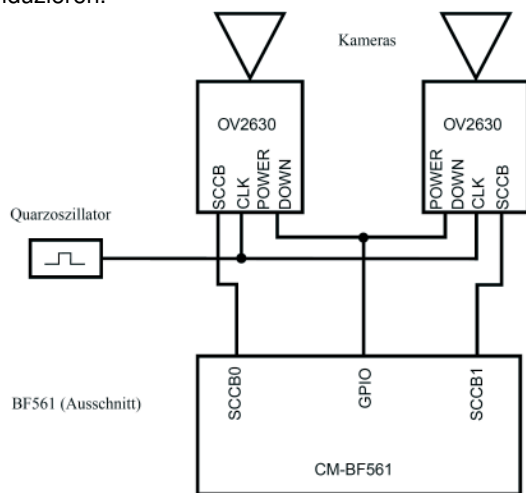


BILD 3. Synchronisation Stereokameras

Die Synchronisation der beiden Kameramodule erfolgt implizit aufgrund der Tatsache, dass es sich bei den Bildsensoren um endliche Automaten handelt. Beide Module (Bild 3) beziehen den erforderlichen Systemtakt aus einem Oszillator. Zusätzlich wird der POWERDOWN-Eingang beider Kameramodule, über den sich die Sensoren deaktivieren lassen, über einen digital-Ausgang des DSP geschaltet. Dies hat zur Folge, dass beide Module gleichzeitig den Betrieb aufnehmen und aufgrund der Versorgung aus einem Oszillator synchron Bilddaten zur Verfügung stellen.

### 3.1.3. Live-Video Schnittstelle

Das VTOL-MAV MD4-200 verfügt über eine integrierten Videotransmitter zur Übertragung von Live-Videodaten an die Bodenstation, der in der Lage ist, Videoinformationen im so genannten FBAS - Format zu übertragen. Dieses Übertragungsformat wird als Basisbandsignal bei der Ausstrahlung von analogem Fernsehen verwendet und hat dadurch eine weitere Verbreitung gefunden.

Da die Bilddaten nach Verarbeitung durch den Prozessor

zunächst digital vorliegen, ist eine Digital-Analog Wandlung unter Berücksichtigung des gewünschten Übertragungsstandards erforderlich. Hierzu wird der integrierte Schaltkreis ADV7393 [4] der Firma Analog Devices eingesetzt. Dieser wandelt digitale Bilddaten in ein analoges PAL oder NTSC-Signal um. Die digitale Ansteuerung des ADV7393 erfolgt über eine 8-Bit breite LVTTTL Schnittstelle im YCrCb-Format, die die zweite Schnittstelle des ADSP-BF561 nutzt. Die Signalisierung des Beginns einer neuen Bildzeile beziehungsweise eines neuen Bildes erfolgt entsprechend der Ansteuerung der beiden Kamerasensoren über die Signale HSYNC und VSYNC. Neben der digitalen Videoschnittstelle findet zur Konfiguration wesentlicher Betriebsparameter eine I2C Schnittstelle Verwendung, die den Zugriff auf die Konfigurationsregister des ADV7393 ermöglicht.

### 3.1.4. Massenspeicher

Neben der Übertragung von Echtzeit-Videodaten ist auch die Möglichkeit, Einzelbilder und Videosequenzen dauerhaft abzuspeichern, eine wesentliche Anforderung an die Systemplattform. Diese ist in Form eines Slots zur Aufnahme von Micro-SD Karten realisiert. Bei diesen Karten handelt es sich um Speicherkarten in Flash-Technologie mit Abmessungen von 11 mm x 15 mm x 0,7 mm und einer Masse von 0,4 Gramm.

Als elektrische Schnittstellen stehen wie bei dem Vorgänger der Micro-SD Karten, der SDCard, sowohl der proprietäre SD Standard als auch die *Serial Peripheral Interconnect* (SPI) Schnittstelle zur Verfügung. Zwingend vorgeschrieben ist allerdings ausschließlich der SD Standard, die Implementierung des SPI Modus ist seitens der Hersteller von Micro-SD Karten optional [5]. Da jedoch alle gängigen Hersteller diesen Übertragungsmodus unterstützen und die Spezifikation im Gegensatz zum SD Standard ohne Lizenzgebühren verfügbar ist, ist die Anbindung des Micro-SD Slots über die SPI-Schnittstelle des ADSP-BF561 realisiert.

### 3.1.5. Layout der Systemplattform

Bild 4 gibt einen Überblick über die relevanten Systemkomponenten sowie die eingesetzten Schnittstellen zur Verbindung der Komponenten; dargestellt sind darüber hinaus die erforderlichen Bandbreiten dieser Schnittstellen, die bei der Entwicklung von Hardwaredreibern und Applikationen zu berücksichtigen sind.

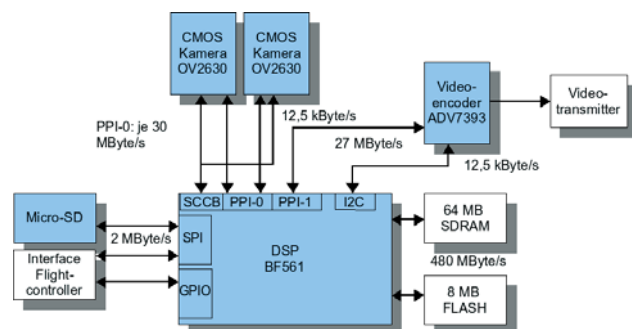


BILD 4. Struktur Systemplattform

Zentrales Element ist das DSP-Modul CM-BF561, das neben einem Analog Devices ADSP-BF561 Blackfin

Prozessor den erforderlichen SDRAM-Arbeitsspeicher, einen NOR-Flash Speicher für das Betriebssystem sowie weitere erforderliche Komponenten wie Spannungs- und Takterzeugung enthält.

Die beiden CMOS-Kameramodule OV2630 sind direkt über die in den ADSP-BF561 integrierte erste PPI-Schnittstelle an den Prozessor angeschlossen. Zur Übertragung von Live-Videodaten wird der Videoencoder ADV7393 über die zweite PPI-Schnittstelle des ADSP-BF561 mit digitalen Videoinformationen versorgt, die nach Erzeugung eines analogen PAL-Signals über einen in die das MAV integrierten Videotransmitter an eine Bodenstation übermittelt werden. Zur Konfiguration der Kameras und des Videoencoders sind jeweils eine SCCB beziehungsweise I2C Schnittstelle erforderlich, die per Softwareemulation über digitale Ein-/Ausgänge (GPIO) realisiert sind.

Zur dauerhaften Speicherung von Einzelbildern oder auch kurzen Videosequenzen ist ein Slot zur Aufnahme von Micro-SD Speicherkarten vorgesehen, die wie deren Vorgänger, die MMC-Karten, über eine SPI-Schnittstelle verfügen. Die Integration der Bildverarbeitungsplattform in die bestehenden Flight- und Navigation-Controller des MAVs erfolgt über den SPI-Bus sowie über GPIO Schnittstellen.

## 3.2. Software

Als Softwarebasis für die entwickelte Systemplattform findet das Betriebssystem  $\mu$ CLinux Anwendung [6], da dieses als quelloffenes System eine Anpassung an die Hardwareplattform und Peripheriekomponenten ermöglicht und gleichzeitig bereits eine große Anzahl wesentlicher Betriebssystemfunktionen integriert.  $\mu$ CLinux ist eine Portierung des Betriebssystems Linux auf Prozessorplattformen ohne Memory Management Unit, zu der auch die Blackfin-Familie gehört.

### 3.2.1. Das Betriebssystem $\mu$ CLinux

Der Quellcode des  $\mu$ CLinux Systems ist mit Ausnahme der hardwarerelevanten Module identisch mit der Portierung für PC-Systeme und ermöglicht somit die Verwendung von Softwarewerkzeugen, Programmbibliotheken und Applikationen, die für PC-basierte Umgebungen unter Linux entwickelt wurden, mit geringem Portierungsaufwand auf einer eingebetteten Plattform wie der des Blackfin. Des Weiteren gestattet dies die Entwicklung und Validierung von Applikationen der automatisierten Bildverarbeitung in einer leistungsfähigen PC-Umgebung, die mit geringem Aufwand auf der Zielplattform implementiert werden können.

Eine lauffähige  $\mu$ CLinux - Umgebung besteht aus den folgenden Komponenten:

- 1)  $\mu$ CLinux -Kernel
- 2) RAM-Disk
- 3) Applikationen

Der Kernel oder Betriebssystemkern stellt dabei die wesentliche Komponente des Systems dar. Er ist für die Initialisierung des Systems, die Verwaltung von Arbeitsspeicher und Prozessen, die Koordination von

Zugriffen der Applikationen auf die Hardware sowie der Überwachung von Zugriffsrechten und Sicherheitsrichtlinien zuständig. Außerdem stellt dieser die für den Zugriff von Applikationen auf Massenspeicher erforderlichen Schnittstellen zur Verfügung.

Die RAM-Disk trägt das zum Betrieb erforderliche Dateisystem und wird im Laufe des Startvorgangs aus einem nichtflüchtigen Speicher wie beispielsweise einer Festplatte, einer CD oder einer Flash-Speicher in den Hauptspeicher geladen. Dieses Dateisystem enthält alle Applikationen und Hilfsprogramme, die für den Betrieb des Systems benötigt werden. Bei PC-Systemen ist dieses Dateisystem üblicherweise auf einer Festplatte abgelegt, dies ist jedoch auf einem eingebetteten System im Allgemeinen aufgrund der beschränkten Rahmenbedingungen nicht möglich.

Zu den Applikationen gehören neben den produktiven Anwendungen, die dem Gesamtsystem die gewünschte Funktion verleihen, auch Softwarewerkzeuge, die für den Start und die Verwaltung des Betriebssystems erforderlich sind.

### 3.2.2. Echtzeitfähigkeit

$\mu$ CLinux unterstützt als Multitasking/Multiuser-Betriebssystem die quasi-parallele Ausführung mehrerer Programme oder *Tasks*. Dies erschwert jedoch die Realisierung von Applikationen unter Echtzeitbedingungen. Echtzeit bedeutet in diesem Kontext, dass eine bestimmte Funktion innerhalb einer definierten Zeit abgeschlossen ist, beziehungsweise dass das System auf eine externe Anforderung innerhalb einer definierten Zeit reagiert. Unter einem Multitasking-Betriebssystem ist jedoch ohne weitere Vorkehrungen nicht vorhersagbar, zu welchem Zeitpunkt eine Applikation Rechenzeit erhält und somit fristgerecht auf eine Anforderung reagieren kann.

Um die Anforderung der Echtzeitfähigkeit ausgewählter Funktionen erfüllen zu können, sind zeitkritische Funktionen wie beispielsweise die Übertragung von Live-Videodaten an den Videoencoder ADV7393 nicht als Applikationen realisiert, sondern als Modul in den Betriebssystemkern integriert. Da Kernelmodule über eine erhöhte Priorität gegenüber Applikationen außerhalb des Betriebssystemkerns verfügen, können diese nicht bei der Ausführung unterbrochen werden beziehungsweise werden bei Bedarf schneller aufgerufen als Applikationen. Des Weiteren entfällt die zeitaufwendige Kommunikation durch Kopieren von Speicherbereichen zwischen Applikationen und Hardwareschnittstellen über das Kernel.

Eine weitere Maßnahme zur Erhöhung der Echtzeitfähigkeit ist die Reduktion des Arbeitsspeichers, der unter der Verwaltung des Kernels steht, dargestellt in Abbildung 5. Dies gestattet einen deterministischen Zugriff auf einen Teil des Arbeitsspeichers durch Applikationen. Der Zugriff auf diesen Speicherbereich erfolgt dabei über die physikalische Adresse des Bereichs; hierbei wirkt sich die nicht vorhandene Memory Management Unit des BF561 positiv aus, da andernfalls ein direkter Zugriff der Applikationen auf den Speicher nicht möglich wäre.



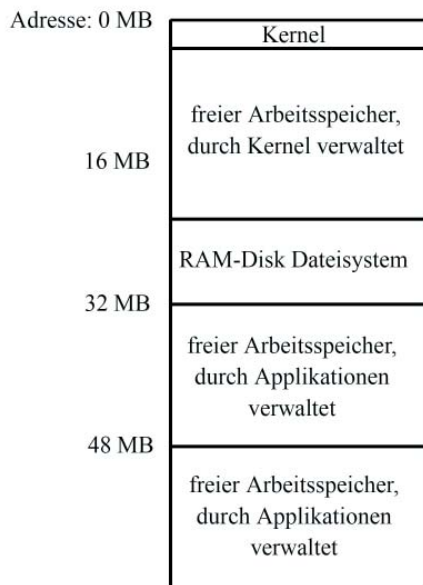


BILD 5. Speicherlayout unter µCLinux

Ein weiterer Vorteil dieser Reduktion ist in den spezifischen Eigenschaften des SDRAM-Speichers begründet. Im Gegensatz zu asynchronem SRAM-Speicher, auf den nach Anlegen der gewünschten Speicheradresse lesend und schreibend zugegriffen werden kann, ist beim Zugriff auf SDRAM-Speicher ein Busprotokoll zu beachten. Dieses gestattet den Zugriff nur nach vorheriger Initialisierung. SDRAM-Speicher besitzt einen linearen Adressraum, der jedoch in Bänke und Zeilen segmentiert ist. Vor Zugriff auf eine Speicherzelle ist zunächst die Bank und Zeile, in der sich die Speicherzelle befindet, zu aktivieren. Diese Aktivierung erfordert mehrere (Speicher-)Taktzyklen, so dass während der Umschaltung auf eine andere Zeile oder Bank keine Zugriffe erfolgen können. SDRAM-Speicher ist typischerweise in 4 Bänke sowie eine von der Speichergröße abhängigen Anzahl Zeilen mit einer Größe von jeweils 2 KByte segmentiert [7]. Pro Bank kann nur jeweils eine Zeile gleichzeitig aktiviert sein. Der SDRAM-Controller des ADSP-BF561 [8] ist in der Lage, in jeder der 4 Bänke eine Zeile dauerhaft geöffnet zu halten, so dass durch ein entsprechendes Layout der Daten im SDRAM-Speicherbereich die Anzahl der erforderlichen Zyklen zur Aktivierung der SDRAM-Zeilen reduziert werden kann. Da der Beginn einer SDRAM-Bank sich an einer durch 16 MB teilbaren Adresse befindet, resultiert die Reduktion des durch das Kernel verwalteten Speicherbereichs sowie die Aufteilung der verbliebenen Speicherbereiche zur Verwaltung durch die Applikationen, wie in Bild 5 dargestellt, in einer verkürzten Zugriffszeit auf Daten im SDRAM-Speicher.

#### 4. ANWENDUNGEN

Zur Demonstration und Validierung der Leistungsfähigkeit sind zwei Applikationen auf der Systemplattform realisiert, die ein breites Spektrum der zur Verfügung stehenden Komponenten nutzen. Bei der Implementierung sind dabei die spezifischen Eigenschaften der DSP-Hardwarearchitektur zu berücksichtigen.

Die wesentliche Herausforderung bei der effizienten Implementierung mathematischer Algorithmen auf modernen Mikroprozessor- und DSP-Plattformen ist derzeit, dem Prozessorkern die zu verarbeitenden Daten in ausreichender Geschwindigkeit zur Verfügung zu stellen. Die Taktfrequenz moderner Prozessoren in der Größenordnung von 1 GHz liegt deutlich über der gegenwärtig existierender Speichertechnologien wie zum Beispiel DDR-SDRAM mit einer Frequenz von 200 MHz.

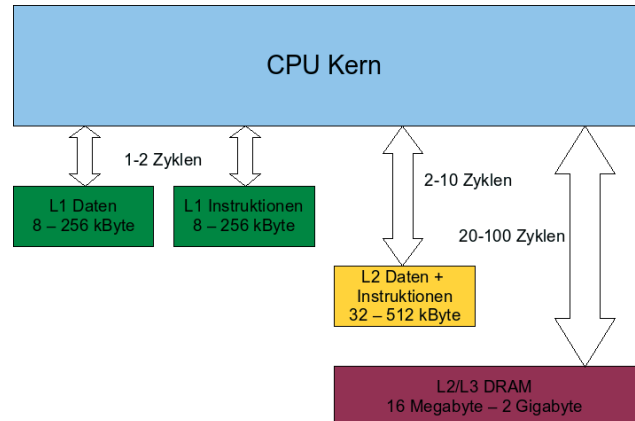


BILD 6. Zugriffsgeschwindigkeit der Speichertypen

Abbildung 6 verdeutlicht die hierarchische Speicherstruktur des ADSP-BF561 und stellt die verschiedenen Geschwindigkeiten und Speichergrößen der einzelnen Speicherbereiche dar.

Bei der Implementierung von Bildverarbeitungsalgorithmen ist aufgrund der erheblichen Unterschiede hinsichtlich der Geschwindigkeit von CPU-Kern und Arbeitsspeicher neben einer effizienten Auslegung des Algorithmus bezüglich mathematischer Optimierungen insbesondere auf das Layout der zu verarbeitenden Daten im Speicher zu achten. So sollten rechenintensive Anwendungen möglichst auf Daten zugreifen, die sich innerhalb des L1-Speichers befinden, da im Allgemeinen Zugriffe auf diesen Bereich mit der Geschwindigkeit des CPU-Kerns erfolgen. Moderne DSP-Implementierungen verfügen zwar über schnelle Zwischenspeicher, so genannte *Caches*, die Zugriffe auf häufig benötigte Daten im externen Speicher beschleunigen, bei großen Datenmengen, wie sie in der automatisierten Bildverarbeitung anzutreffen sind, beziehungsweise hohen Datenraten aber nahezu wirkungslos sind. Dies stellt einen gravierenden Unterschied zu Prozessoren im Umfeld der PC-Architektur dar, die mittlerweile über große Caches verfügen und so in der Lage sind, ohne Optimierungen des Speicherlayouts eine hohe Leistungsfähigkeit zu erbringen.

##### 4.1. Erzeugung und Übertragung stereographischer Bilddaten

Die Wahrnehmung der Tiefeninformation ist ein wesentlicher Bestandteil der menschlichen Fähigkeit, sich in seiner Umgebung zu orientieren. Die Vermittlung der räumlichen Tiefe unter Verwendung einer Kamera ist allerdings aufgrund des zweidimensionalen

Abbildungsverfahrens nicht möglich, da dies dazu führt, dass sich alle dreidimensionalen Informationen der Umgebung innerhalb einer Ebene befinden. Eine weitere Problematik ergibt sich bei der Darstellung künstlich erzeugter stereographischer Bilddaten, da hierbei die physiologischen Besonderheiten der Bildwahrnehmung durch den Menschen berücksichtigt werden müssen. Der Einsatz technischer Systeme zur Verbesserung des Situationsverständnisses seitens des Anwenders durch dreidimensionale Darstellung kann jedoch zur Erhöhung der Leistungsfähigkeit des Gesamtsystems beitragen, da der Anwender durch eine realitätsnahe Darstellung in geringerem Maße mit der Interpretation des Sachverhaltes belastet wird und somit Entscheidungen schneller und mit höherer Zuverlässigkeit getroffen werden können. Da dies auch auf Einsatzszenarien des MAVs MD4-200 wie der Unterstützung von Rettungskräften zutrifft, ist die Realisierung eines derartigen Systems wünschenswert.

#### 4.1.1. Entstehung des Tiefeneindrucks

Prinzipiell basiert die menschliche Bildwahrnehmung auf der Projektion der dreidimensionalen Welt auf die zweidimensionale Oberfläche der Netzhaut und enthält somit ebenfalls keine Information über die räumliche Tiefe der Umgebung. Die Tiefenwahrnehmung wird dadurch rekonstruiert, dass beide Augen bedingt durch einen räumlichen Abstand eine Situation aus einer jeweils unterschiedlichen Perspektive wahrnehmen. Die über die Netzhaut erfassten Bilder werden im Gehirn überlagert und über die geometrischen Beziehungen der relativ zueinander verschobenen Perspektiven eine dreidimensionale Wahrnehmung vermittelt. Die entworfene Systemplattform bildet über die Verwendung zweier Kameramodule, die in einer Ebene mit einem konstanten Abstand fixiert sind, das Verfahren der menschlichen Wahrnehmung nach.

#### 4.1.2. Signalpfad der Bilddaten

Abbildung 7 verdeutlicht den Signalpfad der Bilddaten von der Aufnahme mittels Kamerasensoren bis zur Visualisierung auf einem entsprechenden Ausgabemedium. Nach der Erfassung der Bildinformationen durch die Kameramodule liegen diese in Form von Einzelbildern vor, deren Farbinformation nach dem Bayer-Schema kodiert sind. Die Aufnahme der beiden Perspektiven hat dabei zum selben Zeitpunkt, das heißt synchron, zu erfolgen, da andernfalls ein weiterer räumlicher Versatz zwischen den beiden Perspektiven entsteht, der die Wahrnehmung der Tiefeninformation beeinträchtigt. Die nach dem Bayer-Schema kodierten Pixeldaten sind zunächst werden zunächst so interpoliert, dass jeder Pixel durch die drei Farbkanaäle Rot, Grün und Blau repräsentiert wird, danach erfolgt die Fusion der beiden Einzelaufnahmen so, dass diese in einem Format vorliegen, das für eine Übertragung an den Benutzer geeignet ist und auf einem entsprechenden Ausgabemedium dargestellt werden kann. Vor der Übertragung erfolgt eine Farbraumkonvertierung in das YUV- beziehungsweise YCbCr-Format, um die Bilddaten an die Charakteristik des Übertragungsstandards anzupassen. Zuletzt erfolgt die Übermittlung der digital vorliegenden über eine analoge Videoschnittstelle an die Bodenstation sowie die Darstellung für den Anwender.

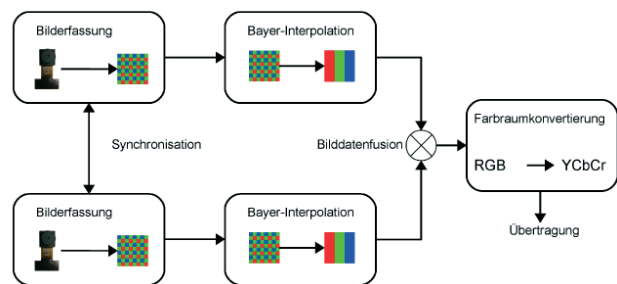


BILD 7. Signalpfad bei Übertragung von Stereobildern

#### 4.1.3. Erzeugung und Übertragung des PAL-Signals

Die Übertragung der Bilddaten erfolgt mittels des bereits in das MAV integrierten Videotransmitters in Form eines FBAS-Signals im PAL-Standard. Dies hat den Vorteil, dass zum einen die Darstellung auf gängigen Videoausstattungen möglich ist, zum anderen die Übertragung der beiden Perspektiven durch Fusion der Bilddaten mit nur einer Übertragungsstrecke möglich ist.

Die Fusion der Bilddaten wird dadurch vorgenommen, dass die linke Kameraperspektive als ungerade Halbbilder, die rechte Kameraperspektive als gerade Halbbilder übertragen werden. Da die Reihenfolge der Halbbilder seitens des Empfängers detektiert werden kann, ist eine Trennung der Perspektiven und separate Darstellung auf einem geeigneten Medium möglich. Diese Übertragung hat als Konsequenz, dass die vertikale Bildauflösung jedes Einzelbildes auf ein Halbbild mit 288 Zeilen beschränkt ist. Die fehlenden Zeilen müssen zur Vermeidung einer verzerrten Darstellung bei der Anzeige interpoliert werden. Das durch die Fusion beider Kameraperspektiven in Form von zwei PAL-Halbbildern erzeugte Videosignal wird in einem Videopuffer im SDRAM-Speicherbereich abgelegt.

Die Übertragung des Videopuffers an die PPI-Schnittstelle zur Übermittlung an den Videoencoder erfolgt unter Zuhilfenahme von *Direct-Memory-Access* (DMA) Transfers. Es ist dabei erforderlich, dass diese Transfers kontinuierlich erfolgen, um einen Verlust der Synchronisation mit dem FBAS-Signal zu verhindern. Dies ist insofern problematisch, als dass die Struktur eines PAL-Bildes die Übertragung in einem einzigen, sich wiederholenden DMA-Transfer unmöglich macht. Ein einzelnes Bild gemäß PAL-Standard gliedert sich in die folgenden vier Komponenten:

- ungerades Halbbild
- gerades Halbbild
- Vertikale Austastlücke des ungeraden Halbbildes
- Vertikale Austastlücke des geraden Halbbildes

Zusätzlich existiert in jeder Zeile eine horizontale Austastlücke, die jedoch als konstanter Offset zu jedem einzelnen Halbbild betrachtet wird und daher keine separate Komponente darstellt. Die Austastlücken sind historisch bedingt und stellen den Zeitraum dar, in dem sich bei analogen Fernsehgeräten der Elektronenstrahl zur nächsten Zeile beziehungsweise an den Anfang des nächsten Halbbildes bewegt.

Die Abfolge der einzelnen Halbbilder ergibt sich aus dem PAL-Standard und ist in Abbildung 8 dargestellt. Zur

Realisierung dieser Abfolge, ohne eine aufwändige und zeitkritische Konfiguration des DMA-Transfers vor jeder PAL-Komponente vornehmen zu müssen, bietet das DMA-System des ADSP-BF561 die Möglichkeit, die Konfiguration der DMA-Transfers automatisch erfolgen zu lassen. Dieses System gestattet die Konfiguration der wesentlichen Parameter eines DMA-Transfers mittels so genannter *Deskriptoren*. Diese Deskriptoren stellen eine Datenstruktur dar, die Parameter wie Start- und Zieladressen, Länge des Transfers oder Wortbreite definieren. Darüber hinaus besteht die Möglichkeit, diese Deskriptoren zu verketteten, das heißt, dass am Ende der Übertragung einer PAL-Komponente automatisch die Informationen über die sich anschließende Komponente aus einem weiteren Deskriptor gelesen werden.

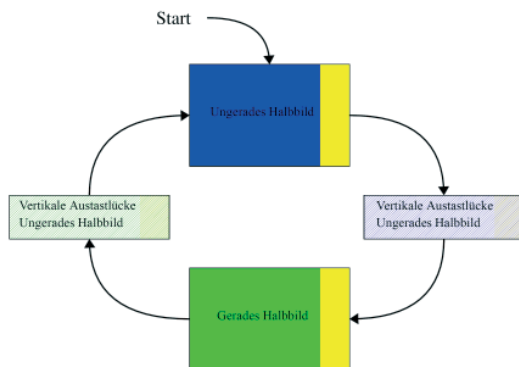


BILD 8. Übertragung einer Bildsequenz mittels PAL-Verfahren

#### 4.1.4. Visualisierung

Zur Darstellung der stereographischen Bilddaten ist es erforderlich, die in beiden Halbbildern kodierten, unterschiedlichen Perspektiven den Augen des Betrachters separat darzustellen. Dazu bietet sich die Verwendung einer Videobrille mit jeweils einem Monitor pro Auge an. Häufig verfügen diese bereits über die Möglichkeit, beide Halbbilder auf den unterschiedlichen Monitoren darzustellen, so dass eine weitere Verarbeitung der Bilddaten nicht erforderlich ist.

## 4.2. Schätzung des optischen Flusses

Unter dem Optischen Fluss wird in der Bildverarbeitung das Vektorfeld verstanden, das die zweidimensionale Bewegungsrichtung eines jeden Bildpunktes in Bildsequenzen definiert. Die Schätzung des optischen Flusses ist ein Standardverfahren der automatisierten Bildverarbeitung und wird beispielsweise bei der Berechnung von Korrespondenzen in der Stereobildverarbeitung, bei der Implementierung von Algorithmen zur Bewegungsschätzung und -kompensation sowie zur Entfernung von Bildverzerrungen angewandt. Die Schätzung des optischen Flusses zählt somit zu den Standardwerkzeugen der modernen Bildverarbeitung.

Ein weit verbreitetes Verfahren zur Bestimmung des Optischen Flusses ist das Verfahren nach Lukas-Canade [9], das in Bezug auf die Komplexität des Algorithmus

effizient ist. Die Effizienz der Implementierung betrifft in erster Linie die Anzahl der erforderlichen Rechenoperation zur Schätzung des Optischen Flusses, der benötigte Speicherplatz ist maßgeblich durch die Auflösung der zu verarbeitenden Bilddaten bestimmt und wird durch das Verfahren nach Lukas-Canade nicht verringert.

Des Weiteren erfordert der Algorithmus Operationen über einen Speicherbereich hinweg, der mindestens der Größe der Rohdaten entspricht. Diese Eigenschaften stellen bei der Implementierung auf einer modernen PC-Architektur keine Einschränkungen dar, da PC-Prozessoren über ausreichend große CPU-Caches und komplexe Speichersysteme zur Optimierung der Zugriffe verfügen. Diese Bedingungen sind bei digitalen Signalprozessoren nicht gegeben, so verfügt der ADSP-BF561 über lediglich 100 KByte Cache pro CPU-Core, eine Größe, die von den Applikationen und dem Bildmaterial in der digitalen Bildverarbeitung bei weitem überschritten wird. Außerdem ist die Architektur der Blackfin Familie auf eine sequentielle Datenverarbeitung optimiert. Die wesentlichen Einschränkungen der Leistungsfähigkeit des Lukas-Canade Algorithmus auf einer DSP-Plattform sind daher die Operationen über einen großen Speicherbereich hinweg, da die Vorzüge der DSP-Architektur durch die im Verhältnis zu der Verarbeitungsgeschwindigkeit des CPU-Kerns hohen Latenzzeiten beim Zugriff auf externen SDRAM-Speicher zunichte gemacht werden. Die Cache-Speicher des ADSP-BF561 sind zu klein, um die wahlfreien, nicht-sequentiellen Zugriffe des Lukas-Canade Algorithmus beschleunigen zu können.

#### 4.2.1. Census Transformation

Der implementierte Algorithmus basiert auf der Census-Transformation [10]. Die Grundüberlegung der Nutzung der Census-Transformation zur Schätzung des Optischen Flusses ist in Abbildung 9 dargestellt. Erfolgt in einer Bildsequenz durch eine Bewegung der abgebildeten Objekte eine Änderung der Pixelwerte, so bleiben diese innerhalb einer begrenzten Umgebung konstant oder ändern sich, beispielsweise durch Beleuchtungsänderungen, relativ zu einander in nur geringem Umfang. Diese einzelnen Umgebungen lassen sich in Bildsequenzen durch die Eigenschaft lokalisieren, dass die Farb- oder Helligkeitswerte auch nach Bewegung innerhalb eines begrenzten Bereichs relativ zueinander konstant bleiben. Der Optische Fluss lässt sich folglich auch dadurch schätzen, dass einzelne Pixel in Bildsequenzen über die Ähnlichkeit ihrer Umgebung verfolgt werden. Der zu ermittelnde Verschiebungsvektor ergibt sich dabei aus der Differenz der Positionen einer derartigen Umgebung in aufeinanderfolgenden Bildern. Der Optische Fluss wird aus diesem Vektor durch die Verknüpfung mit der Dauer des Abtastintervalls der Bildsequenz berechnet.

Zur Ermittlung der Ähnlichkeit von definierten Ausschnitten oder Umgebungen eines Bildes ist eine Charakterisierung dieser Bereiche mit relativ zueinander konstanten Pixeln erforderlich. Diese Charakterisierung erfolgt im Rahmen der Census-Transformation durch die Klassifikation der Helligkeit der umgebenden Pixel in Bezug auf einen zentralen Referenzpixel.

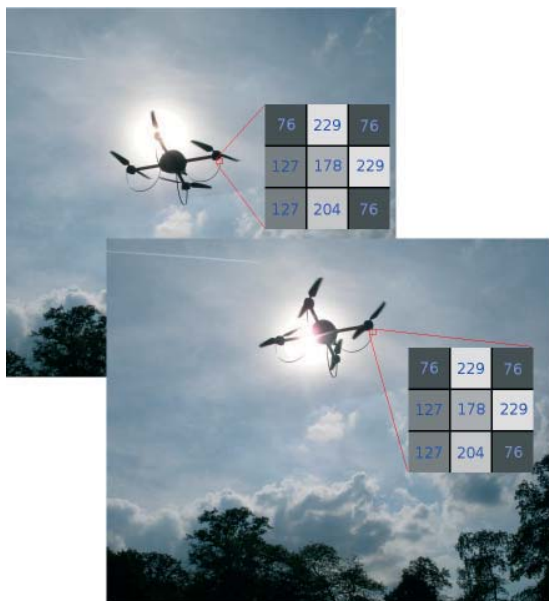


BILD 9. Prinzip Census-Transformation: invariante Pixelhaufen

Gemäß Abbildung 10 erfolgt diese Klassifikation nach den Kategorien *heller*, *dunkler* oder *identische Helligkeit*. Es ist nicht erforderlich, dass ein Umgebungspixel den exakten Grau- oder Farbwert des zentralen Pixel besitzt, um in die Kategorie *identische Helligkeit* zu fallen, vielmehr wird bei der Klassifikation ein Toleranzband definiert, so dass alle Pixel, deren Wert sich innerhalb dieses Toleranzbandes befindet, als Pixel mit identischer Helligkeit klassifiziert werden. Dies soll durch Rauschen verursachte Effekte reduzieren.

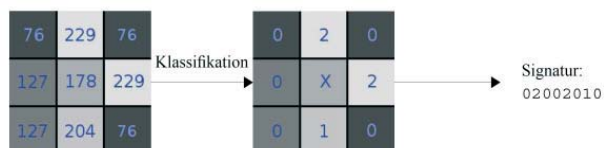


BILD 10. Erzeugung einer Pixelsignatur durch Klassifikation nach Helligkeit

Die relative Helligkeit kann pro umgebendem Pixel nach der Klassifikation gemäß den benannten Kategorien als ein Zahlenwert von [0..2] dargestellt werden, durch Verknüpfung aller klassifizierten Nachbarpixel zu einem Vektor lässt sich somit eine eindeutige Charakterisierung für den zentralen Pixel finden. Dieser Vektor kann in eine Ganzzahl umgewandelt werden, indem die einzelnen Klassifikationen mit einem Zahlenwert von [0..2] als Stellen einer Zahl im Trinärsystem interpretiert werden. Die so gefundene Repräsentation der Umgebung eines Pixel wird auch als Signatur oder *Patch* bezeichnet. Sind zwei Patches identisch, bedeutet dies, dass auch die Umgebung zweier Pixel relativ zu diesen Pixeln bezüglich der Farb- oder Helligkeitswerte identisch ist. Mit den obigen Grundüberlegungen zu der Bewegung von Objekten in Bildsequenzen lassen sich somit Verschiebungsvektoren und daraus der Optische Fluss ermitteln.

Durch die Census-Signaturen ist eine Identifizierung unterschiedlicher Pixelumgebungen mit einer hohen Sicherheit möglich, dies resultiert in einer zuverlässigen Detektion der Verschiebungsvektoren. Gleichzeitig fällt dabei eine große Anzahl an Patches an, die aufgrund der Struktur der umgebenden Pixel, beispielsweise bei Umgebungen mit einer homogenen Helligkeit, keine Informationen zur Ermittlung der Verschiebungsvektoren enthalten. Zur Eliminierung derartiger Signaturen werden Filter eingesetzt, die irrelevante Patches anhand verschiedener Kriterien entfernen. Zu diesen Kriterien gehört neben der erwähnten homogenen Umgebung auch die Anzahl identischer Patches. Überschreitet diese Anzahl eine bestimmte Schwelle, lassen sich die zusammengehörige Umgebungen nicht mehr eindeutig identifizieren und gestatten somit keine eindeutige Zuordnung.

Abbildung 11 verdeutlicht das Verfahren zur Schätzung des Optischen Flusses mittels der Census-Transformation: Zunächst werden für alle Pixel eines Grauwertbildes die Census-Signaturen ermittelt und in einer Tabelle abgespeichert. Dabei werden mit einem ersten Filter irrelevante Patches entfernt. Die Liste mit irrelevanten Patches muss dabei a priori bekannt sein. Sie enthält beispielsweise Patches, die eine homogene Umgebung beschreiben. Im nächsten Schritt werden für das folgende Bild der Sequenz ebenfalls die Signaturen

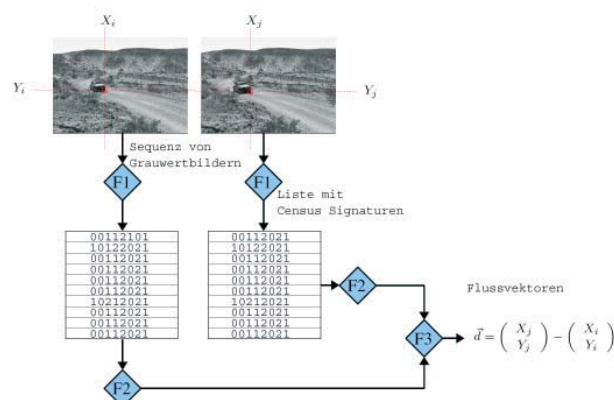


BILD 11. Schätzung des optischen Flusses mittels Census-Transformation

aller Pixel berechnet und gefiltert in einer zweiten Tabelle abgespeichert. Die in den beiden Tabellen enthaltenen Signaturwerte mittels eines zweiten Filters auf Patches ohne relevanten Informationsgehalt eliminiert. Dieses zweite Filter wendet dabei Informationen an, die zur Laufzeit des Algorithmus ermittelt werden, insbesondere findet hierbei das Kriterium der maximalen Anzahl identischer Signaturen Anwendung. Die Zuordnung der Patches zu den Komponenten des zu ermittelnden Verschiebungsvektors erfolgt über ein drittes Filter, über das Randbedingungen wie Beleuchtungsänderungen oder räumliche Distanz der potentiellen Signaturen zueinander einfließen.

Der wesentliche Vorteil der Schätzung des Optischen Flusses mittels Census-Transformation in Bezug auf die Implementierung auf einer DSP-Plattform liegt in der Form der Speicherezugriffe. Die Berechnung der Census-Signaturen erfordert einmalig eine sequentielle Verarbeitung aller Pixeldaten eines Grauwertbildes. Bedingt durch die Verwendung des modularen



Filterkonzepts werden irrelevante Patches bereits bei der Berechnung der Signaturen eliminiert und so der Speicherbedarf erheblich dezimiert. Bereits nach dem ersten Filterschritt existieren lediglich  $10^4$ - $10^5$  Signaturen, die bei der Weiterverarbeitung berücksichtigt werden müssen. Die Liste dieser Signaturen lässt sich innerhalb der Speicherbereiche einer CPU ablegen, auf die mit einer geringen Latenzzeit zugegriffen werden kann. Alle weiteren Filteroperationen auf diese Liste können effizient und mit hoher Geschwindigkeit durchgeführt werden, da der erforderliche Speicheraufwand erheblich geringer als bei Verwendung des Lukas-Canade Algorithmus ist.

#### 4.2.2. Implementierung der Census-Transformation

Um die Leistungsfähigkeit der Schätzung des optischen Flusses mittels Census Transformation beurteilen zu können, wird dieser Algorithmus auf der realisierten Systemplattform implementiert. Es zeigt sich dabei (Tabelle 3), dass die Verwendung der Census-Transformation auf der Systemplattform bei einer Frequenz von 3 Hz eine ausreichende Anzahl an Flussvektoren ermitteln kann, um eine zuverlässige Schätzung des Optischen Flusses zu ermöglichen.

Bildauflösung	720 Pixel x 240 Zeilen
Ermittelte Flussvektoren	970
Frequenz der Schätzung	3 Hz

TAB 3. Schätzung des optischen Flusses mittels Census-Transformation: Leistungsdaten

## 5. ZUSAMMENFASSUNG

Es wird in diesem Beitrag die Implementierung einer Systemplattform für die Echtzeit-Bildverarbeitung an Bord eines VTOL-MAVs vorgestellt. Wesentliche Randbedingungen sowie Entwurfsziele werden skizziert und daraus die Auswahl der erforderlichen Hardwarekomponenten abgeleitet. Die Integration der verschiedenen Hardwarekomponenten wird erläutert und die Umsetzung einer Softwareumgebung für den Betrieb der Hardwareplattform dargelegt. Als wesentlicher Bestandteil der Arbeit werden die Unterschiede zu einer PC-basierten Plattform herausgearbeitet und Besonderheiten bei der Portierung von bestehenden Applikationen aufgezeigt.

Masse	44,0 g
Abmessungen	10,0 cm x 6,5 cm x 1,1 cm
Versorgungsspannung	5 – 17 Volt DC
Stromaufnahme	400 mA bei 5 Volt
Kameraauflösung	2 x 1600 Pixel x 1200 Zeilen
Bildfrequenz	variabel, bis 30 Hz

TAB 4. Technische Daten Systemplattform

Es zeigt sich dabei, dass trotz erheblicher Einschränkungen hinsichtlich Abmessungen, Masse und Leistungsaufnahme Anwendungen der modernen Bildverarbeitung auf der fliegenden Plattform eines elektrisch betriebenen Quadropters realisiert werden können. Die definierten Randbedingungen können gemäß Tabelle 4 eingehalten werden.

Mittels der Realisierung einer Applikation zur Erzeugung

und Übertragung von stereographischen Bilddaten sowie der Implementierung eines Algorithmus zur Schätzung des Optischen Flusses wird die Praxistauglichkeit der Systemplattform demonstriert.



BILD 12. Systemplattform – Hardware

Zukünftige Arbeiten sind bei der Integration der Systemplattform in die bestehende Avionik angesiedelt, dazu ist zunächst ein Protokoll für die Kommunikation mit Flight- und Navigation-Controller zu entwerfen. Weitere Schritte können in der Unterstützung einer automatischen Landung mittels Bildverarbeitung oder der Stützung der Navigationslösung durch Abgleich mit Kartendaten liegen.

## 6. LITERATUR

- [1] Delfly MAV. – <http://www.delfly.nl>
- [2] Microdrones Produktflyer MD4-200. – microdrones GmbH, Siegener Str. 227, 57223, Kreuztal
- [3] Omnivision Technologies, Inc.: *Datenblatt OV2630 Color CMOS UXGA (2.0 Megapixel) CameraChip*. Version 1.3. March 2005
- [4] ANALOG DEVICES: *Datenblatt ADV7390/ADV7391/ADV7392/ADV7393 Low Power, Chip Scale 10-Bit SDHD Video Encoder*. One Technology Way, P.O. Box 9106, Norwood, MA: Analog Devices Inc., 2006
- [5] SD Card Association: *SD Specifications, Part 1 Physical Layer, Simplified Specification*. Version 2.00. 2006
- [6] Blackfin Linux Project. – <http://blackfin.uclinux.org>
- [7] MICRON TECHNOLOGY: *Datenblatt: Synchronous DRAM MT48LC128M4A2, MT48LC64M8A2, MT48LC32M16A2*. 8000 South Federal Way, Post Office Box 6, 83707 Boise: Micron Technology, Inc., 2000
- [8] ANALOG DEVICES: *Blackfin BF561 Hardware Reference Manual*. One Technology Way, P.O. Box 9106, Norwood, MA: Analog Devices Inc., 2005
- [9] BRUCE D. LUCAS, Takeo C.: *An Iterative Image*

Registration Technique with an Application to Stereo Vision. In: *Proceedings of Imaging Understanding Workshops*, 1981

[10] STEIN, Fridtjof: Efficient Computation In: *Proceeding of DAGM*, 2004