



Projektzusammenfassung zum Vorhaben
„Entwicklung einer Multilayercodierung“

im Rahmen des Eurostars Projekts
E! 10301 ManiCode
**„Effiziente und echtzeitfähige Erstellung und Codierung von
mit überlagerten Inhalten erweiterten TV-Übertragungen“**

Holger Meuel, Jörn Ostermann

Leibniz Universität Hannover
Institut für Informationsverarbeitung
Appelstr. 9A
30167 Hannover

1 Vorhabensziele

Im Rahmen von ManiCode sollte ein Videoanalyse und -synthesystem für die automatische Erstellung und effiziente Codierung von lokalisierten TV-Videodatenströmen erstellt werden. Dabei werden Teile realer Videodatenströme durch künstliche Inhalte ersetzt. Beispielsweise soll Werbung in Livesportübertragungen durch regionalisierte Werbung ersetzt werden. Im Projektteil *Entwicklung einer Multilayercodierung* sollte eine Videocodierung entwickelt werden, welche die Übertragung mehrerer lokalisierter Videodatenströme in einem Transportdatenstrom vereint und somit die Übertragung über nur einen Satellitenkanal ermöglicht. Bisherige Systeme übertragen lokalisierte Videodatenströme separat in jeweils eigenen, teuren Satellitenkanälen. Zur Projektmitte sollte ein Softwaredemonstrator als Meilenstein präsentiert werden. Zum Projektende sollte ein funktionstüchtiger Hardwaredemonstrator erstellt werden. Um einen zeitnahen Markteintritt zu ermöglichen, sollte am Ende des Projekts eine gründliche Evaluation des System durchgeführt werden.

2 Vorhabensinhalte und -ergebnisse

2.1 Multilayercodiersystem

Für eine effiziente Codierung über das Maß des standardisierten Videokompressionsverfahrens AVC hinaus, werden die mit künstlichen Inhalten überlagerten Videosignale (Localized Feeds, LFs) nicht wie sonst üblich codiert, sondern es werden in einem Vorverarbeitungsschritt zunächst in allen Localized Feeds jeweils mit dem Originalvideo (World Feed, WF) gemeinsame Bildbereiche durch monochrome Bildbereiche ersetzt (grün in Abbildung 1 (f)-(h)). Bildbereiche, die überlagerte Inhalte enthalten, werden unangetastet gelassen. Der WF und die vorverarbeiteten LFs werden nun von parallel laufenden AVC-Videoencodern codiert. Die großflächigen monochromen Bereiche können sehr effizient codiert werden, weshalb statt der typischerweise anfallenden Datenrate des World Feeds von 20 Mbit/s jeder Localized Feed bei gleicher Qualität nur weniger als 2 Mbit/s benötigt. Nach der eigentlichen Videodecodierung müssen die im Vorverarbeitungsschritt in den LFs modifizierten Bildbereiche durch einfaches Kopieren der entsprechenden Bildbereiche aus dem World Feed wieder korrekt rekonstruiert werden. Welche Bildbereiche zwischen einem jeweiligen Localized Feed und dem World Feed unterschiedlich sind, wird durch pixelweise Berechnung am Encoder ermittelt. Pro $16 \times 16 \text{ pel}^2$ großem Bildblock wird mittels eines Bits signalisiert, ob ein Block überlagerte Inhalte enthält oder nicht. Diese zusätzliche Information wird in den Videodatenstrom als private *Supplementary Enhancement Information* (SEI) eingebettet und kann vom Decoder zur korrekten Rekonstruktion der Localized Feeds wieder extrahiert werden. Alle Videodatenströme (World Feed + beliebige Anzahl Localized Feeds) wurden in einen MPEG-Transport-Strom (MPEG-TS) gemultiplext und können somit auch standardkonform über etablierte Kanäle wie DVB-Systeme oder per TCP/IP-Übertragung über das Internet verbreitet werden.

2.2 Encoderoptimierungen

Oben beschriebene Realisierung mit vorverarbeiteten Localized Feeds funktioniert bezüglich der Codiereffizienz hervorragend. Da jedoch dennoch komplette Full-HD-Bilder codiert werden

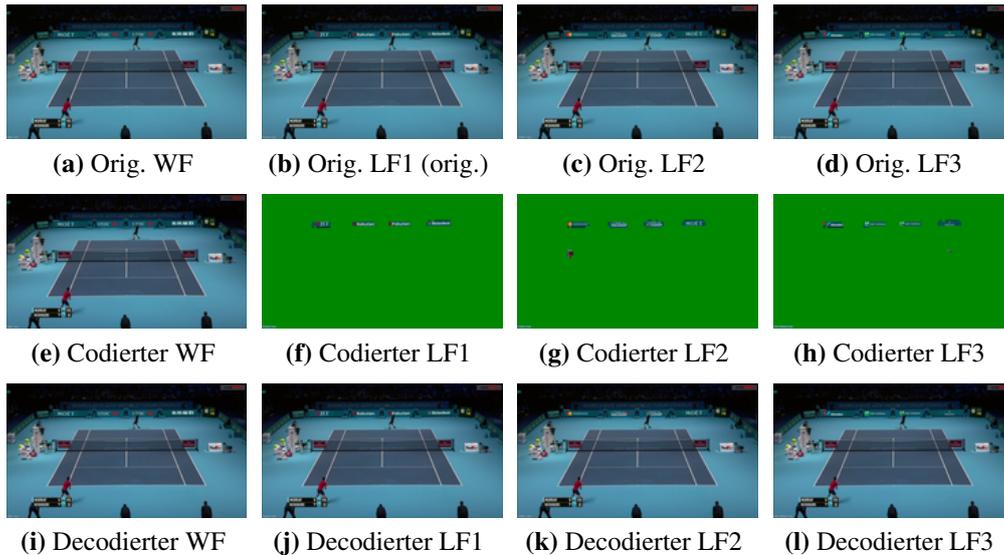


Abbildung 1: Original, codierte sowie decodierte und rekonstruierte Bilder des World Feeds (WF) und der 3 Localized Feeds (LFs).

müssen, ist die Codiergeschwindigkeit ähnlich zu der des World Feeds, obwohl *eigentlich* nur kleine Bildbereiche eines jeden LFs tatsächlich nutzbaren Inhalt enthalten. Daher wurden Verfahren zur Geschwindigkeitsoptimierung des Encoders erforscht und untersucht. Es wurde eine modifizierte Videoencodersteuerung realisiert, welche alle mit dem World Feed gemeinsamen Bildbereiche eines jeden LFs zwangsweise auf den sogenannten Skip-Mode setzt. Im Skipmode werden kein Prädiktionsfehler oder Bewegungsvektoren übertragen, welche üblicherweise einen Großteil der Datenrate benötigen, so dass dieser Mode sehr effizient ist. Die im erzwungenen Skipmode codierten Bildbereiche enthalten nun ebenso wie in der oben beschriebenen Version des Softwaredemonstrators (APs 5 und 6) keine korrekten Inhalte mehr und müssen decoderseitig wie oben beschrieben in einem Nachverarbeitungsschritt rekonstruiert werden. Durch die modifizierte Steuerung entfällt die encoderinterne, extrem zeitintensive Bewegungsschätzung für alle gemeinsamen Bereiche (grüne Flächen in Abbildung 1). Dadurch kann die Codiergeschwindigkeit deutlich erhöht werden. Die benötigte Encodierzeit skaliert ungefähr linear mit der prozentualen Bildfläche. Für eine Sportszenario, bei welchem üblicherweise weniger als 5 % der Bildfläche von künstlich überlagerten Inhalten bedeckt werden, ist die Zeitersparnis im Vergleich zur Codierung des gesamten Bildes also 95 %. Auch für eine Hardwarerealisation ist eine Beschränkung der Bewegungsschätzung auf bestimmte, dem Encoder bereits bekannte Bildbereiche sinnvoll.

3 Nutzen bzw. Anwendungsmöglichkeiten der Ergebnisse

Die Ergebnisse des Projekts sollen vom Projektpartner uniqFEED kommerzialisiert werden. Verhandlungen zur Lizenzierung der effizienten Multilayercodierung laufen zwischen uniqFEED und dem TNT. Eine Erfolgsabschätzung ergab ein hohes Marktpotential für auf den Projektergebnissen beruhende Produkte.