

Analys av H.265/HEVC (inte färdig)

Kim Le, 36950

Kandidatavhandling i Datateknik

Handledare: Jan Westerholm

Fakulteten för naturvetenskaper och teknik

Åbo Akademi

2016

Innehåll

Innehåll	2
Referat	3
1 Inledning	4
2 Videokodning	5
2.1 Bakgrund	5
2.2 ITU-T	6
2.3 ISO/IEC MPEG	6
3 Videokodningsprocessen	7
4 Utvecklingen av videokodningsstandarder	8
4.1 Blockbaserad hybridkodning	9
4.2 H.261	10
4.3 H.262/MPEG-2 Video	11
4.5 H.264/AVC	11
5 H.265/HEVC	12
5.1 Parallellbearbetning	13
6 Avslutning	13
7 Referenser	14

Referat

I dagsläget är TV-apparater med stöd för UHD (Ultra högupplösning, eng. *Ultra High Definition*) allt vanligare på marknaden, vilket betyder att upplösningen fyrdubblats från den största möjliga upplösningen för HDTV (High-Definition television). Detta ger konsumenterna en mycket skarp och detaljrik bild. För att möjliggöra sändning av video i en sådan upplösning som ligger inom UHD-spektrat krävs videokomprimeringsstandarder som till exempel VP9 och High Efficiency Video Coding (HEVC). Denna avhandling kommer endast att sätta fokus på HEVC-standarderna. Skillnader mellan HEVC och föregående standarder kommer att behandlas samt förbättringar och metoder som gör standarderna så högeffektiva.

Nyckelord: HEVC, High Efficiency Video Coding, videokodning

1 Inledning

High Efficiency Video Coding (HEVC) även känt som H.265, är en videokodningsstandard som utvecklades av två standardiseringsorgan, ITU-T Video Coding Experts Group (VCEG) och ISO/IEC Moving Picture Experts Group (MPEG). Utvecklingen av HEVC görs i form av ett samarbete mellan dessa två organ som tillsammans gått ihop som Joint Collaborative Team on Video Coding (JCT-VC). Arbetet på HEVC började 2010 och den första versionen av projektet färdigställdes i april 2013[1, 2]. HEVC är inte JCT-VC:s första projekt, utan de ligger även bakom föregående standarder som H.262/MPEG-2 Video och H.264/MPEG-4 AVC (Advanced Video Coding), varav den senare fortfarande är i bruk, till exempel hos tv-operatörer i Sverige för sändning av SVT:s (Sveriges Television) högupplösta kanaler [3], men eftersom fler tv-apparater med stöd för ännu högre upplösningar än HDTV (High-Definition television) produceras, har en efterfrågan på en effektivare standard uppstått.

En annan orsak till behovet av en modernare videokodningsstandard är den kraftiga ökningen av internet-tv. till exempel Netflix som är en webbplats för beställvideo med ett brett sortiment av filmer och serier. Cisco, ett ledande företag inom nät- och internetlösningar uppskattar att all videorelaterad internettrafik av konsumenter, på en global nivå, kommer att utgöra från 80-90 procent av all internettrafik genererad av konsumenter år 2019. Detta skulle utgöra en kraftig ökning från 64 procent år 2014. Enligt Cisco var 59 procent av beställvideoströmningen högupplöst, år 2014. De uppskattar även här en ökning, till 70 procent år 2019 [4].

Som tidigare nämnts har marknaden börjat blicka mot framtiden, det vill säga en uppgång för UHD (Ultra High-Definition), främst 4K UHD [5]. Med detta menas att upplösningen består av minst 8 miljoner aktiva pixlar, där minimigränsen för antalet pixlar på den horisontella axeln är 3840 och 2160 på den vertikala axeln [5]. Uppspelning av video i UHD-formatet betyder att mängden data ökar och detta kan bli ett problem för mobila apparater som smarttelefoner och pekplattor, speciellt då största delen av internettrafik förväntas att komma från mobila apparater. Cisco uppskattar att endast 33 procent av all internettrafik förväntas komma från

persondatorer 2019 [3]. HEVC utvecklades därför med målet att halvera överföringshastigheten. Syftet med denna avhandling är att ge en god överblick över HEVC samt att beskriva metoderna som gör HEVC till en modernare och förbättrad standard.

2 Videokodning

Videokodning är processen som behandlar video i form av komprimering, det vill säga komprimering och dekomprimering. Komprimering betyder att datafilens lagringsutrymme förminskas, ibland på kvalitetsens bekostnad. Den omvända processen kallas för dekomprimering. Kodning kan uppdelas i två olika sorter, förstörande (eng. *lossy*) eller icke-förstörande komprimering (eng. *lossless*). Förstörande komprimering innebär en reduktion i storleken på datan liksom en reduktion av kvaliteten på datan. I icke-förstörande komprimering sker ingen förlust av datakvalitet, det vill säga att den ursprungliga datan kan rekonstrueras, men storleken på datan förminskas inte i samma grad som i förstörande komprimering [6].

Videokodning är ytterst viktigt eftersom digital video i sin ursprungliga, okomprimerade form tar väldigt mycket lagringsutrymme. Detta kan göra sändning av digital video omöjlig. Europeiska radio- och TV-unionen (European Broadcasting Union, EBU) rekommenderar alla sina medlemsländer att sända digital tv i formatet 720p/50. Detta format innebär en hög upplösning på 1920 pixlar horisontellt och 720 pixlar vertikalt samt att bilden visas med hastigheten 50 bilder per sekund. Formatet har en överföringshastighet på 1,485 Gbit/s för den okomprimerade signalen, men alla tv-apparater har inte mottagare som klarar av en signal av denna magnitud [3]. Dessa problem är aktuella och varför utveckling av effektivare videokodningsstandarder är ytterst viktig.

2.1 Bakgrund

Eftersom världen i dagsläget blivit ett kommunikationssamhälle är det ännu viktigare än förr med internationella standarder. Standarder ger producenter en bra grund för

att vara konkurrenskraftiga, men de kan även gynna konsumenterna, till exempel i form av hårdvarukostnader. Standarder existerar även för att säkerställa kompatibilitet mellan system.

2.2 ITU-T

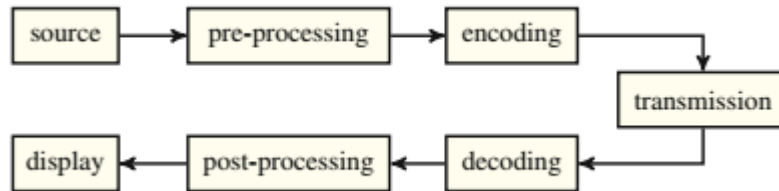
Internationella teleunionen (ITU) grundades 1865 som International Telegraph Union, efter att 20 europeiska stater undertecknat fördraget vid den första Telegrafkonventionen. Målet med unionen var att tillsammans överkomma geografiska problem som telegrafmeddelanden medförde. I och med teknologiska framgångar, till exempel telefonen och radion, bestämde unionen sig för ett namnbyte vid en konferens 1932. Namnbytet trädde i kraft 1934 och namnet används ännu idag. År 1947 erkändes ITU som ett av FN:s fackorgan. År 1993 omorganiserades ITU och kommittén CCITT som ansvarade för telefon- och telegrafkommunikation ersattes med ITU:s Telecommunication Standardization Sector (ITU-T) [7]. I dagsläget sker utvecklingen av visuell kodning under ITU-T Study Group 16 som även är kända som Video Coding Experts Group (VCEG). VCEG har även bidragit till andra innovationer som bildkomprimeringsstandarden JPEG [8].

2.3 ISO/IEC MPEG

Moving Picture Coding Experts Group (MPEG) bildades 1988 av Leonardo Chiariglione och gruppens uppgift var att utveckla standarder för kodning av både visuella data och ljuddata. Gruppen arbetar under ISO/IEC JTC 1 (International Organization for Standardization/International Electrotechnical Commission Joint Technical Committee 1), som ansvarar för standardisering inom it [9]. MPEG har framställt viktiga standarder som MPEG-1 och MPEG-2, varav den senare har använts som standard för digital television.

3 Videokodningsprocessen

Processen för videokodning för ett generellt videokodningssystem består av sju steg, videoanskaffning, prebearbetning, kodning, överföring, avkodning, postbearbetning och uppspelning. Varje steg av processen kommer att presenteras kortfattat.



Figur 1 Ett blockdiagram för ett generellt videokodningssystem [10].

I videoanskaffningssteget skall videosekvensen matas ut i digital form. En videosekvens är en serie av bilder med ett visst tidsintervall mellan varje bild. Videosekvensens källa kan till exempel vara analoga inspelningar och datoranimerat material. Källan har olika egenskaper beroende på sitt ursprung, som signalkomponenter som interfererar med datainnehållet. Rörelseoskärpa eller artefakter som kvarstår från en digitaliserad analog källa är vanliga signalkomponenter.

Vid prebearbetning utförs olika operationer som trimning och omvandling av färgformat på den inmatade videosekvensen. Huvuduppgiften är dock att ta bort olika oönskade delar, till exempel artefakter som nämndes i föregående stycke. I detta steg skall även bildens egenskaper som upplösningen och färgformatet behandlas. Specifik bearbetning av materialet för videokodningssyften är inte ovanligt. Detta leder till att kodaren i nästa steg komprimerar videon på ett effektivare sätt.

Kodaren omvandlar videosekvensen som prebearbetats till en kodad representation av en bitström. Detta görs för att möta de krav som målapplikationen ställer. Bitströmmen som genereras innehåller information om bland annat kodningsmetoder som används och den överföringshastighet som kodningen medför till nästa skede. Kodaren inkapslar sedan bitströmmen till ett transportformat som förberedelse för de kommande stegen.

I överföringsskedet levereras videomaterialet som behandlats av kodaren till mottagarsidan. Det är ytterst viktigt att systemet är designat på ett sådant sätt att det uppnår sina ändamål. Ett exempel är paketorienterade system där förvrängningar av signalen kan hindra att en felfri överföring lyckas. För att systemet skall möta kraven implementeras därför metoder som skyddar mot paketförlust och som kan återställa förlorade paket.

Då bitströmmen anländer till avkodarskedet så lagrar avkodaren den mottagna bitströmmen i en buffert. Samtidigt börjar avkodaren att rekonstruera den komprimerade datan till en videosekvens med de instruktioner kodaren angivit. Avkodaren lagrar utöver den mottagna bitströmmen, avkodade bilder i en skild buffert. Dessa bilder kan användas som referens vid avkodning. I stycket ovan nämndes paketförluster vid överföring. Ibland kan inte datan som gått förlorad under överföring återställas. Avkodaren har metoder som försöker återställa så mycket som möjligt av sådan förlorad data. Dessa metoder kallas feldöljande (eng. *error concealment*).

I postbearbetningsskedet utförs operationer på den avkodade videosekvensen, till exempel bildförbättring. Om videosekvensens format inte uppnår kraven för uppspelning utförs operationer på bildegenskaperna precis som under prebearbetningen.

Det sista steget är uppspelningsfasen. Nyare standarder har ofta metoder för att använda videomaterialets egenskaper till hjälp för att skapa en bättre visuell upplevelse. Detta kallas för användbar videoinformation (eng. *video usability information*). Den användbara informationen packas in i bitströmmen i ett tidigare skede och avkodaren kan sedan rekonstruera videon med de rätta uppspelningsparametrarna [10].

4 Utvecklingen av videokodningsstandarder

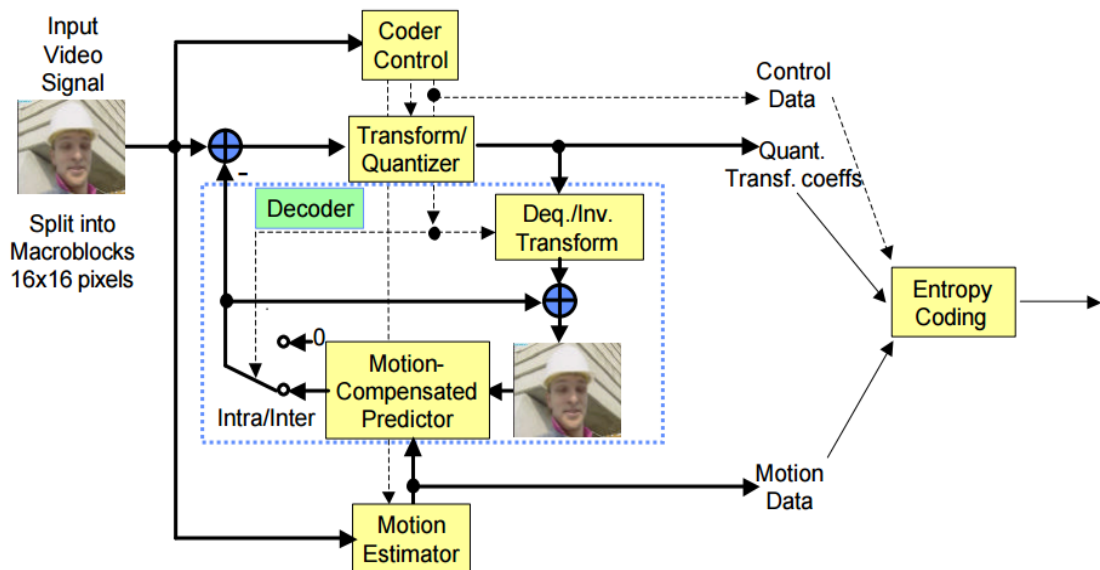
Detta kapitel kommer kortfattat att presentera föregående standarder och deras kodningsverktyg som med varje generation av standarder bidrog till effektivare kodning. Däremot kommer H.120 som var den första standarden för digital video

inte tas upp eftersom dess struktur är så olik HEVC och de andra blockbaserade hybridkodningsstandarderna.

4.1 Blockbaserad hybridkodning

Blockbaserad hybridkodning är ett tillvägagångssätt inom videokodning som först togs i bruk i och med standarden H.261. Alla videokodningsstandarder sedan H.261 använder de koncept som den blockbaserade hybridkodningen bidrog med i någon form.

När en videosekvens matas in delas varje bild in i block. Ett block består av 8x8 lumapixlar. Alla block genomgår antingen intrabildskodning eller interbildskodning. Intrabildskodning betyder att kodningen utförs utan någon referens till andra bilder inom samma sekvens, prediktion sker alltså endast mellan bildens blockregioner. Första bilden använder sig därför alltid av intrabildskodning. Skillnader mellan de följande bilderna inom samma sekvens är små, och beror för det mesta på rörelser. Om dessa rörelser väljs för prediktion, så kan kodaren välja rörelsevektorer (eng. *motion vector*) som indicerar hur bildernas block rör sig i förhållande till referensbilden. Detta är interbildskodning och är ett effektivt sätt att hantera redundans i videosekvenser. Med hjälp av rörelsekomensation (eng. *motion compensation*) genereras en prediktionssignal. Signalen innehåller data om blockförskjutningar mellan en blockregion i en färdigt kodad bild och det nuvarande prediktionsblocket. Båda kodningssätten kan generera en felsignal och de potentiella felsignalerna transformeras med linjära avbildningar, till exempel diskret cosinustransform (DCT). Transformationen medför specifika transformationskoefficienter som kvantiseras och efteråt tillsammans med rörelsedata entropikodas [10, 11].

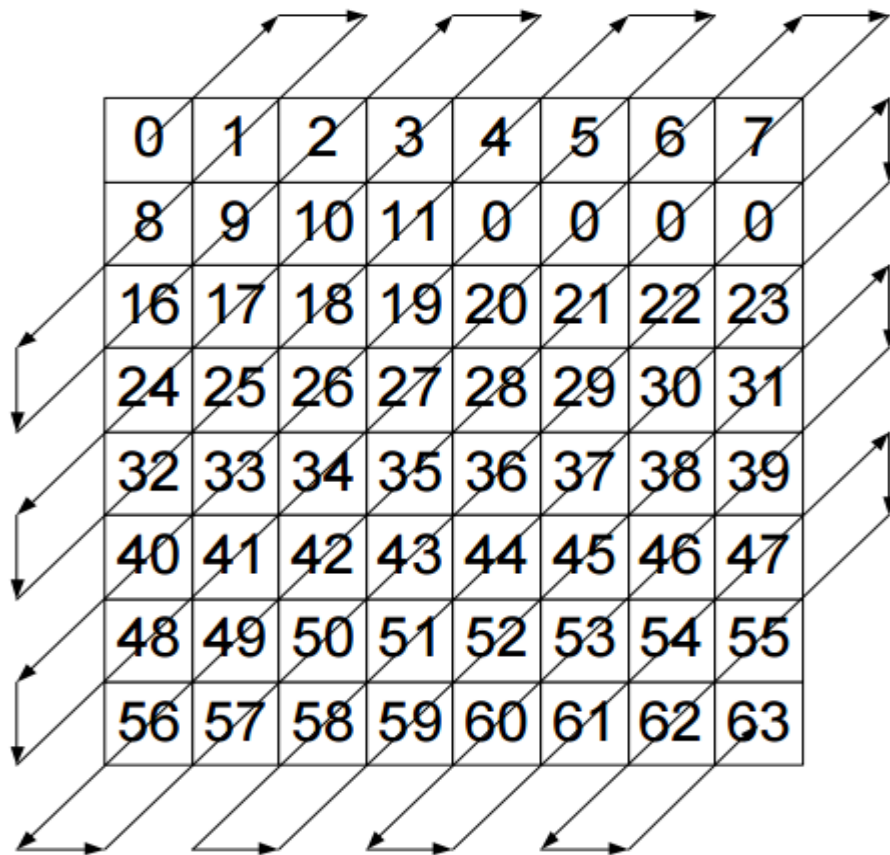


Figur 2 Schema för ett blockbaserat hybridkodningssystem, i detta fall H.261 [12].

4.2 H.261

Standarden H.261 utvecklades av ITU-T och färdigställdes 1990. I delkapitlet ovan nämndes att H.261 var den första videokodningsstandarden som använde sig av blockbaserad hybridkodning, det vill säga att denna standard är grunden till den moderna videokodningen. För denna standard introduceras makroblocket, en 2×2 matris bestående av fyra block, det vill säga ett 16×16 lumblock och ett antal chromablock, beroende på vilket format för chromasampling som används. Detta möjliggjorde rörelsekomensation med 16×16 makroblock. Utöver detta används tidigare kända koncept inom videokodning som DCT med en 8×8 matris, skalär

kvantisering, sicksackavläsning med mera [12].



Figur 3 Sicksackavläsning

TODO: Förklara kort DCT, skalär kvantisering, sicksackavläsning

4.3 H.262/MPEG-2 Video

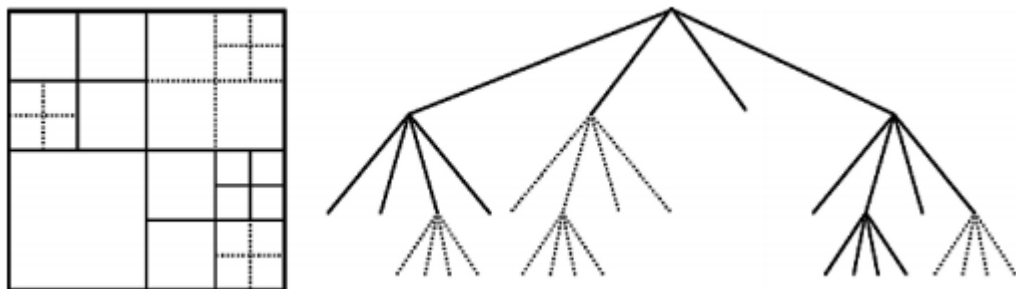
4.5 H.264/AVC

5 H.265/HEVC

High Efficiency Video Coding (HEVC) utvecklades av JCT-VC, som är en sammanslagning av experter inom området från organisationerna ITU-T VCEG och ISO/IEC MPEG. Arbetet började 2010 och den första versionen av standarden färdigställdes april 2013. Med tiden har ett par nya versioner kommit, varav den senaste blev färdig februari 2015. De nya versionerna tillför nya tillägg som till exempel 3D-HEVC, som effektiviserar kodning av avancerade 3D-teckenrutor [2].

Målet med HEVC var en markant förbättring av komprimeringsprestanda gentemot tidigare standarder, eftersom efterfrågan på upplösningar bortom HD ökat. JCT-VC siktade därför på att halvera överföringshastigheten för en likvärdig, perceptuell videokvalitet. De ville i huvudsak fokusera på effektiv kodning vid höga upplösningar som 4Kx2K och bättre parallellbearbetning.

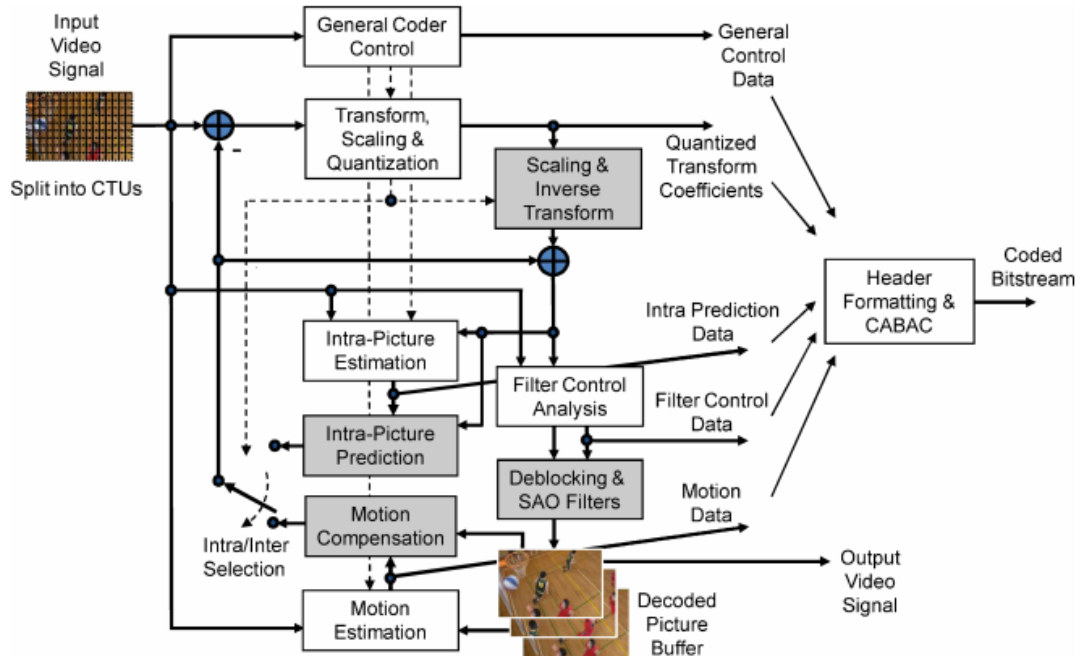
HEVC använder sig av blockbaserad hybridkodning som beskrevs i förra kapitlet. Makroblocken som var bearbetningsenheten i tidigare standarder har i HEVC ersatts med coding tree units (CTU), vars kapacitet kan vara större än makroblockets. En inmatad bild delas in i coding tree blocks (CTB). En CTU innehåller en luma-CTB, två chroma-CTB och deras syntax. Dessa lumablock kan vara av storlek 16x16, 32x32 eller 64x64 lumapixlar, ju större blockstorlek desto bättre komprimering. CTB kan i sin tur delas upp i ännu mindre block, som coding blocks (CB), prediktionsblock (PB) och transformationsblock (TB). Kodaren har tillgång till CTU-syntax, vilket möjliggör uppdelning av CTB i mindre block som nämndes ovan. Vid uppdelningen kan de partitionerade blocken även representeras i form av ett quadträd (eng. *quadtree*). Figur 4 visar hur ett uppdelat block kan se ut [1].



Figur 4 De punkterade linjerna anger TB och de streckade linjerna CB [1].

TODO: Förklara vidare om block, prediktion och så vidare.

Nedan i figur 5 visas ett blockdiagram på en HEVC-kodare.



Figur 5 En HEVC-kodare, de gråa blocken utgör avkodarens metoder [1].

5.1 Parallellbearbetning

TODO: Förklara tiles och wavefront parallel processing (WPP).

6 Avslutning

7 Referenser

- [1] **G.J. Sullivan, J-R Ohm, W-J Han och T. Wiegand.** ”Overview of the High Efficiency Video Coding (HEVC) Standard”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, sidorna 1649-1668, december 2012.
- [2] **Fraunhofer Heinrich Hertz Institute**, [Online]. Tillgänglig: <https://hevc.hhi.fraunhofer.de/> [Hämtad 3 mars 2016].
- [3] **Sveriges Television**, ”Expertinformation”, [Online]. Tillgänglig: <http://www.svt.se/omsvt/om-sandningarna/hdtv/expertinformation> [Hämtad 3 mars 2016].
- [4] **Cisco**, “Cisco Visual Networking Index: Forecast and Methodology, 2014-2019 White Paper”, uppdaterad 27 maj 2015, [Online]. Tillgänglig: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html [Hämtad 3 mars 2016].
- [5] **CTA**, “4K Ultra High-Definition TV”, [Online]. Tillgänglig: <https://www.cta.tech/Consumer-Info/4KUltraHD.aspx> [Hämtad 3 mars 2016].
- [6] **Nationalencyklopedin**, ”Kodek”, [Online]. Tillgänglig: <http://www.ne.se/uppslagsverk/encyklopedi/lång/kodek> [Hämtad 18 mars 2016].
- [7] **ITU**, “Overview of ITU’s History”, [Online]. Tillgänglig: <http://www.itu.int/en/history/Pages/ITUsHistory-page-4.aspx> [Hämtad 3 mars 2016].
- [8] **ITU-T**, “ITU-T SG 16 Standardization on visual coding – the Video Coding Experts Group (VCEG)”, uppdaterad 2013, [Online]. Tillgänglig: <http://www.itu.int/en/ITU-T/studygroups/2013-2016/16/Pages/video/vceg.aspx> [Hämtad 3 mars 2016].
- [9] **MPEG**, ”MPEG Basics”, [Online]. Tillgänglig: <http://mpeg.chiariglione.org/mpeg-basics> [Hämtad 3 mars 2016]
- [10] **M. Wien.** *High Efficiency Video Coding*, Springer-Verlag Berlin Heidelberg. ISBN: 978-3-662-44276-0. Sidorna 23-71, 2015.

[11] **J-R Ohm, G.J. Sullivan, H. Schwarz, T.K Tan och T. Wiegand.**

”Comparison of the Coding Efficiency of Video Coding Standards – Including High Efficiency Video Coding (HEVC)”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, sidorna 1669-1684, december 2012.

[12] **G.J. Sullivan.** “Overview of the International Video Coding Standards (preceding H.264/AVC), uppdaterad juli 2005. [Online]. Tillgänglig:

http://www.itu.int/ITU-T/worksem/vica/docs/presentations/S0_P2_Sullivan.pdf

[Hämtad 30 mars 2016].