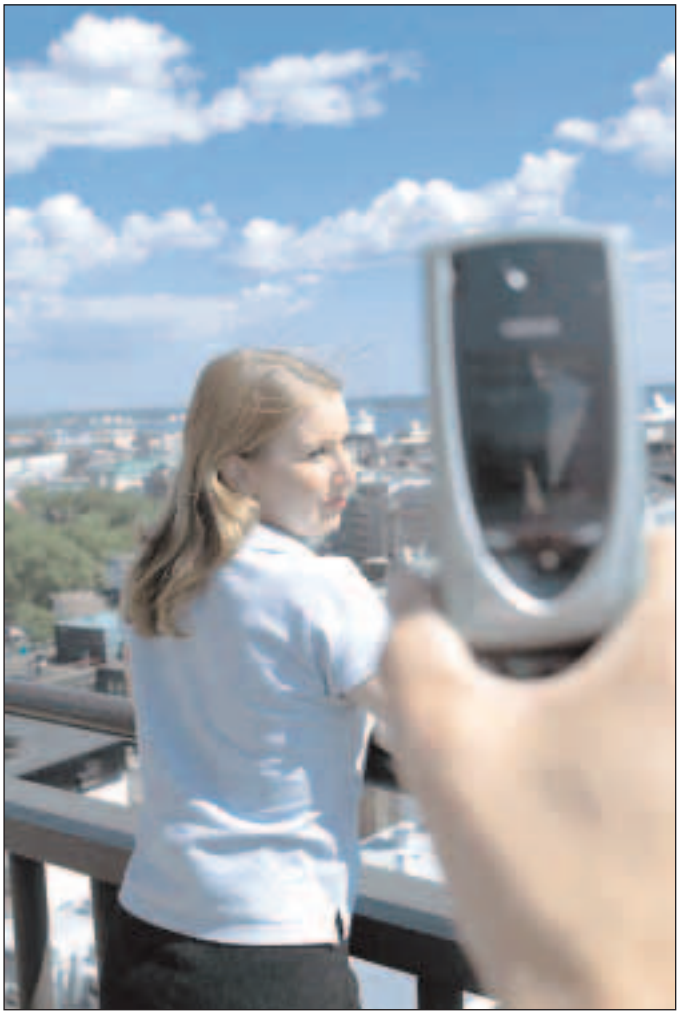
Ensiapua videoviidakossa

Digitaalisen videon helppokäyttöisyys luo nopeasti videoviidakon. Kuinka videot voidaan tallentaa niin, että kiinnostavat videoleikkeet löydetään myöhemmin valtavasta videomateriaalin määrästä? Kuinka löytyneet videoleikkeet siirretään erilaisten verkkojen ja liikkuvien päätelaitteiden välillä?

Johannes Peltola, Marko Jurvansuu, Marko Palola, Jari Korva, Sari Järvinen, Antti Heikkinen, Anna Sachinopoulou, Johan Plomp, Jarkko Nisula, Timo Hotti, Jukka Rantala



Videomateriaalin tuottaminen on helppoa. Sen suurimmaksi ongelmaksi muodostuu valtava määrä. Videon siirtämisen ja tallennuksen lisäksi videoiden selaaminen ja etsiminen on eräs tulevaisuuden haasteista, sillä tallennettu video on käyttökelpoista vasta sitten, kun videon haku- ja siirtotekniikat toimivat saumattomasti yhteen.

Tulevaisuudessa videoarkistosta on mahdollista hakea tietyt videoleikkeet entistä helpommin. Lisä- eli metatietoa kerätään käsin tai automaattisesti esimerkiksi videossa esiintyvistä ihmisistä, asioista, puheesta, videointipaikasta ja -ajasta. Muokkaustyökalut päätelaitteessa antavat mahdollisuuden metatiedon keräämiseen jo videon luontivaiheessa. Videoarkiston ja siihen liittyvien toimintojen hajauttaminen päätelaitteen ja palvelimen välillä helpottaa käyttötilanteeseen oleellisesti liittyvää materiaalin käsittelyä.

Videon hakutekniikat ja halutun video-osan siirtäminen matkapuhelimiin ja muihin langattomiin laitteisiin ovat tärkeitä tutkimuskohteita. Videonsiirtotapoja on kaksi: video siirretään joko ennen katselua tai samanaikaisesti sen kanssa. Jälkimmäinen menetelmä jakaantuu reaaliaikaiseen ja lyhyen viiveen sallivaan. Nämä ovat videon siirron kuumimpia aiheita mahdollistaessaan suorat lähetykset ja suurikokoisten videoviestien katselun aloittamisen ennen kuin koko viesti on siirretty päätelaitteeseen.

Videon siirto matkapuhelimeen vaatii tasaisen yhteyden, jossa verkkopakettien viiveet ovat siedettäviä ja ennustettavia. Lisäksi videon siirtoteknologian pitää arvioida käytetyn yhteyden laatua ja samalla säätää videon tarvitsemaa tiedonsiirtokapasiteettia parantamalla tai heikentämällä kuvan ja äänen laatua.

Hakukoneet videoille

Videomateriaalin hakupalvelut voidaan rinnastaa Internet-hakukoneiden, kuten Googlen, käyttäjille tarjoamiin palveluihin. Automaattisten sisältöanalyysimenetelmien avulla voidaan helpottaa metatiedon tuottamista, sillä liitettävä tieto voidaan koostaa tietokoneavusteisesti tai jopa osittain automaattisesti. Automaattisia kuvaan ja ääneen perustuvia analyysimenetelmiä on kehitetty useiden vuosien ajan lähinnä teollisuuden laaduntarkkailun erikoissovelluksiin liittyen.

Vaikka tarkoin rajatuissa tehtävissä saavutetaankin hyviä tuloksia, on analyysi, jolla kyetään etsimään ja tunnistamaan kaikki erityyppisissä videoissa esiintyvät objektit, vielä kaukana tulevaisuudessa. Täydellistä kuvausta videoiden sisällöstä ei siten vielä saada, mutta eri analyysimenetelmillä voidaan saada viitteitä tallenteen sisällöstä ja käyttää näitä viitteitä hakukoneessa, jolloin jo nykyisillä analyysimenetelmillä voidaan tehostaa hakua.

Esimerkiksi ääniraidan analyysillä video voidaan jakaa osiin siten, että musiikkia, taustaääniä ja puhetta sisältävät kohdat on erotettu toisistaan. Tarkemmalla analyysillä on mahdollista tunnistaa ja luokitella musiikkikappa-

Video jungle

The ease of creating and viewing video content has grown a multimedia jungle. Solutions are needed to store this multimedia material in a manner that facilitates retrieving multimedia presentations, or parts of videos that are of interest to the user. Also delivery of these videos to the user via a range of different networks and terminals needs scaleable and adaptive solutions. VTT, Solid and Hantro in cooperation with their international partners in the Candela project are working on solutions related to video and audio analysis, compression, storage, retrieval and networked delivery.

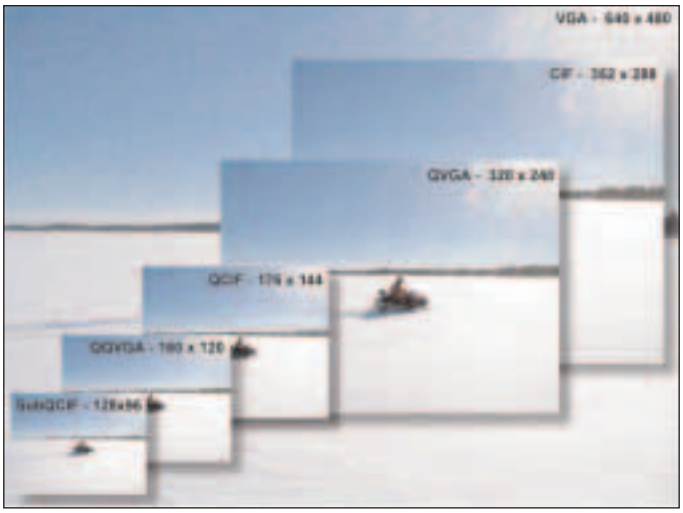
The use of metadata with videos allows the enhancement of videos with information about its contents, the situation the video was made, and other relevant information. This allows efficient retrieval of videos, or their parts, in retrieval tasks. A particularly interesting format for the metadata is MPEG7. Metadata can be added manually, but video and audio analysis techniques can aid by automatically providing part of the metadata. Also techniques for context awareness can contribute to automatic metadata creation. Ontologies can be used to structure the terms used in the metadata and make searches more versatile. The search of video patches may, in addition to the keywords given by the user, also use knowledge about the user's preferences or context may be used to limit the search.

Multimedia data, including the metadata, can be stored in relational databases. This aids in the management and retrieval of videos.

Video delivery over limited networks is a challenging task. Scaleable formats allow for delivery via different networks and to various terminals. Scaleable solutions may e.g. utilise a base layer that contains a low-resolution version of the video, and an enhancement layer that can be used to achieve full quality.

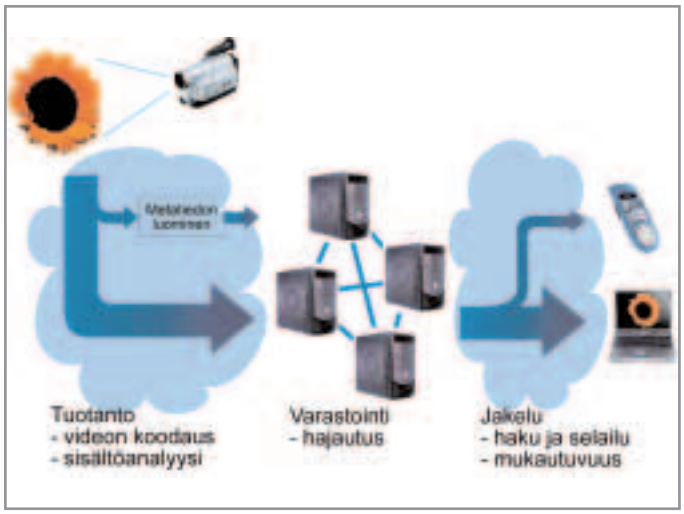
The project aims to build a complete system as a demonstrator for the concepts. These technologies will allow for new services for personal video creation, management and scalable distribution of content.

CANDELA is a EUREKA/ITEA framework project and part of NETS program, which is founded by TEKES. The role of the country representative is fulfilled by VTT's project manager Johannes Peltola (Johannes.Peltola@vtt.fi).



Tyypilliset kuvakoot videosovelluksissa.

Typical image resolutions in video applications.



CANDELA-projektissa toteutetaan videopalvelujen kehittämiseen soveltuva järjestelmä, joka toteuttaa koko käyttöketjun tarjoten alustan videomateriaalin nauhoitukseen, tallennukseen, sisältöanalyysiin ja toimittamiseen eri verkkoja käyttäen erityyppisiin päätelaitteisiin.

The CANDELA project aims to build a complete system for developing new video services. The platform will allow personal video creation and management as well as scalable distribution of video content into different mobile terminals.

leet, eri puhujat sekä eri tyyppinen taustamelu. Näin ääneen perustuvan analyysin avulla voidaan hakea itse tuotetuista videoista esimerkiksi ne, joissa puhuu tietty henkilö. Myös reaaliaikaisen videon pakkausprosessia voidaan hyödyntää esimerkiksi liiketiedon ja kohtausten vaihtumisen ilmaisuun.

Tietoa tiedosta

Suosittu metatiedon määritelmä on ”tietoa tiedosta”. Metatieto voi olla termejä, tekstiä tai kuvia, jotka kuvaavat tieto-olioita ja niiden välisiä suhteita. Tieto-olio voi olla video, kuva, dokumentti tai digitaaliobjekti. Tieto-olio voi esimerkiksi olla video lomasta kreikkalaisella saarella. Metatietoa lomavideosta ovat videon luontiaika ja päivämäärä, paikka ja tieto videossa näkyvistä ihmisistä, asioista, tekstistä, puheesta sekä muista äänistä. Oleellista metatietoa ovat myös videon tekniset yksityiskohdat, kuten kuvan resoluutio, käytetty linssi ja valoisuusaste.

Kun käyttäjä haluaa katsoa videoitaan, hakukone etsii metatiedosta sopivia tuloksia käyttäjän antamien ehtojen mukaan ja palauttaa listan löydetyistä videoista. Käyttäjä valitsee oikean videon ja lopuksi video siirretään hänen päätelaitteeseensa. Videomateriaaliin liittyvän metatiedon sijainti ei riipu videomateriaalin sijainnista, se voi olla tallennettu samaan paikkaan kuin kyseiset tieto-oliot tai eri paikkaan, esimerkiksi metatietopalvelimelle.

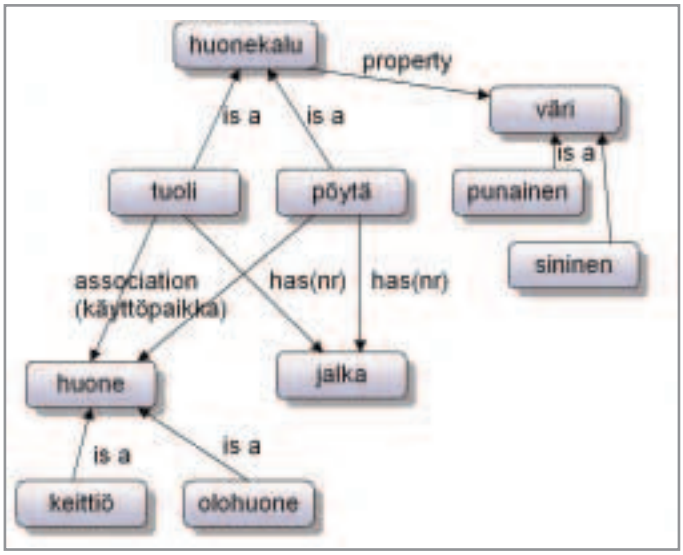
Metatietoformaatteja on monia ja ne ovat yleensä projekti- tai sovelluskohtaisia. On olemassa yksinkertaisia sanastomaisia formaatteja (esimerkiksi Dublin Core), missä käytetään ennalta määrättyjä attribuutteja sisältäviä termejä sekä monimutkaisempia formaatteja, missä voi olla kuvattuna myös olioiden ja termien välisiä suhteita.

MPEG-7 standardoimaan tiedon kuvausta

Yleinen esimerkki tällä hetkellä käytetystä sisällönkuvausmenetelmästä on XML-pohjainen MPEG-7 formaatti. Sen tavoite on mahdollistaa multimediasisällön haku, suodatus ja katselu riippumatta sovelluksista, koodauksesta tai tallennusformaattista. Se tukee sekä reaaliaikaisen että eireaaliaikaisen metatiedon synnyttämistä, sekä push- ja pull-palveluita.

MPEG-7-standardi sisältää säännöt ja työkalut metatiedon tuottamiseen. Metatiedon rooli videohauissa on tärkeämpi kuin tekstihauissa, koska täsmällistä kuvailua (exact match paradigm) ei voida soveltaa digitaaliseen mediamateriaaliin. Uudet videontuottamistyökalut sekä videoanalyysimetodit tuottavat metatietoa, josta on hyötyä käsiteltäessä audiovisuaalisia tieto-olioita.

Metatieto voi olla mediatyyppikohtainen, mediaprosessikohtainen tai sisältökohtainen. Mediatyyppikohtainen metatieto liittyy videoon, ääneen ja muihin mahdollisiin mediatyyppeihin. Esimerkkejä tästä ovat kuvan tekstuuri ja äänen spektrit. Mediaprosessikohtainen metatieto liittyy mediatyypin käsittelyyn, miten video- tai audiomateriaalia voidaan siirtää verkkoon. Sisältökohtainen, semanttinen metatieto taas kuvaa tieto-olion sisällön.



Ontologian avulla hakutermit ja tietoon liitetty metatieto saadaan yhdistettyä ottamalla huomioon asioiden väliset riippuvuudet. Jos käyttäjä hakee esimerkiksi videoita, joissa esiintyy 'huonekalu', saadaan ontologian avulla hakutuloksiksi videot, joihin on merkitty 'tuoli' tai 'pöytä'.

Ontologies can be used to structure the terms used in the metadata and make searches more versatile. For example if user wants to search videos containing furniture, the ontology allows finding videos containing chair or table.

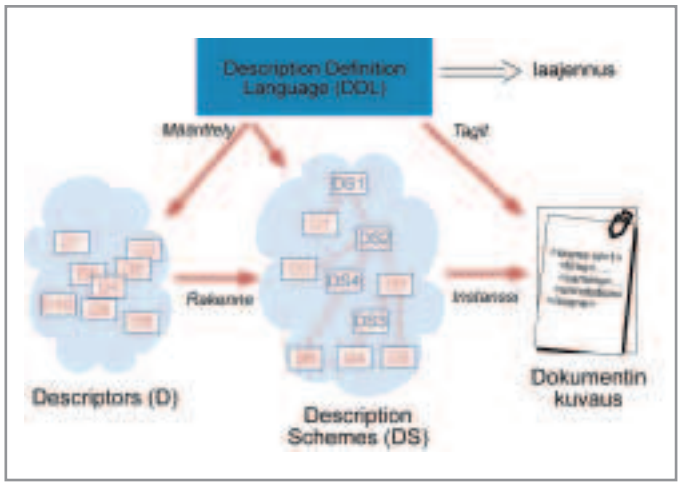
Ontologialla hakutermit kuntoon

Älykkäitä videohakuja käytettäessä käyttäjä voi antaa vapaasti hakutermejä tai valita sopivat ennalta määrätyistä vaihtoehdoista. Haku toimii moitteettomasti vain, jos käyttäjä onnistuu antamaan hakutermeiksi metatietoon tallennettuja termejä. Tällöin hakua voidaan parantaa ontologioiden avulla.

Ontologia on selkeästi määritelty malli, jolla voidaan esitellä ja kuvata oliot, käsitteistö ja niiden väliset suhteet tietyssä tietoalassa. Ontologiahaut ovat tarkoituksenmukaisempia kuin perinteiset täsmähaut, koska ontologialla ei etsitä vain tarkkoja vastaavuuksia. Esimerkiksi käyttäjän etsiessä videomateriaalia, jonka metatiedossa esiintyy termi baari, ontologiahaku sisällyttää tuloksiin myös termit ravintola ja kapakka, koska niiden välillä on selvä yhteys.

Ontologiat voivat myös rajata hakua antamalla käyttäjälle mahdollisuuden valita eri kategorioiden välillä. Käyttäjän avainsanalle jalka ontologia voisi antaa seuraavat kategoriat: huonekalut, eläimet, lääketiede ja kosmetiikka. Lisäksi ontologian avulla voidaan rajoittaa haun tuloksia määrittelemällä monimutkaisia ehtoja kuten ”punainen huonekalu, jolla on neljä jalkaa ja jota käytetään keittiössä”. Tuloksena löytyvät esimerkiksi kaikki punaiset keittiön pöydät ja tuolit, jotka esiintyvät tietokannassa.

Ontologian määritteleminen on



MPEG-7 standardin avulla voidaan kuvata dokumenttien sisältöä. DDL mahdollistaa kuvausten (D) ja kuvausskeemojen (DS) määrittelyn. DDL:n avulla on mahdollista luoda sovelluskohtaisia metatiedon kuvauskenttiä ja määritellä niiden suhteet.

MPEG-7 standard allows describing the content of digital documents. The DDL allows the definition of the MPEG-7 description tools, both Descriptors and Description Schemes. The DDL also allows the extension for specific applications of particular DSs.

Videosoitin näyttää tietoa videon ajallisesta sisällöstä, jolloin videoiden selaaminen on käyttäjälle helpompaa. Kuvassa on jaettu noin puolen tunnin nuoriso-ohjelman tallenne tummiin puhe- ja vaaleisiin musiikkiosiin, jolloin käyttäjän on helppo etsiä tallenteessa olevat musiikkikappaleet tai haastattelukohdat.

Videoplayer is capable of visualising the content of the video for facilitating the efficient browsing of the video clip. 30 minutes long youth TV-program can be seen in dark speech and light music segments. Color view makes it easy for the user to listen which music pieces are located in video or search interview sections.

monimutkainen ja työläs vaihe. Usein ontologia määritelläänkin hyvin suppeaan aihepiiriin liittyen. Kuitenkin useissa käynnissä olevissa projekteissa eri maissa luodaan uusia, laajoja ontologioita. Sopivan ratkaisun ja tarkkuustason löytäminen onkin monesti haasteellista.

Käyttäjän haluamaa aineistoa

Kiinnostavan videomateriaalin etsimisen helpottamiseksi voidaan käyttää rajoituksena myös taustatietoa, esimerkiksi tietoa henkilön mieltymyksistä ja tilanteesta (kontekstista). Näin saadaan aikaan personoitu ja tilannetietoinen palvelu.

Personointiin tarvitaan malli käyttäjän mieltymyksistä eli profiili. Yksinkertaisimmillaan käyttäjä itse määrittelee kiinnostuksensa kohteet. Sitä varten hän voi täyttää avainsanoista koostuvan profiilin kiinnostuksensa ilmaisevilla arvosanoilla. Profiilin täyttäminen on kuitenkin työlästä ja moni käyttäjä jättää sen tekemättä. Siksi on pyritty löytämään automaattisia menetelmiä, joiden avulla hankitaan profiilitietoa.

Automaattiset järjestelmät keräävät yleensä tietoa käyttäjästä huomaamattomasti seuraamalla hänen käyttäytymistään. Valitettavasti järjestelmä pystyy tekemään ehdotuksia vasta seurattuaan käyttäjää riittävän pitkään. Tarvittavaa opetustietomäärä on niin huomattava, etteivät tällaiset järjestelmät ole käytännössä toimineet kovin hyvin. Sen sijaan on pyritty käyttämään apuna muilta samankaltaisilta käyttäjiltä saatua tietoa ehdotusten laatimiseen. Näin esimerkiksi Amazon-verkkokirjakauppa ehdottaa muita vaihtoehtoja kiinnostuksesi mukaisen kirjan perusteella. Videoetsintäpalvelu voisi toimia samalla tavalla, kun käyttäjien määrä on suuri.

Tilannetietoisuus liittyy videomateriaalin tapauksessa kahteen vaiheeseen. Videon luontivaiheessa, varsinkin kun kyse on kuluttajan ottamasta videosta, voidaan videon metatietoon sisällyttää tilannetietoa. Aika on itsestään selvä tilannetieto, ja myös paikka antaa hyvin arvokasta lisätietoa. Paikkatieto saadaan tulevaisuudessa yhä paremmin suoraan selville GPS-, GSM- tai muiden paikannusteknologioiden avulla. Ajan ja paikan lisäksi myös muut tilannetiedot tulevat kyseeseen - lämpötila, valoisuus, paikan luokitus (sisällä, ulkona, autossa...) ja läsnäolevat henkilöt.

Videon luontivaiheeseen liittyvää tilannetietoa voidaan käyttää haun rajaukseen. Toinen vaihtoehto on käyttää hakutilanteeseen liittyvää tilannetietoa videon etsinnässä. Käyttäjän silloinen tilanne otetaan haun rajoitukseksi. Esimerkiksi käyttäjä voi haluta videomateriaalia, joka liittyy hänen fyysiseen sijaintiinsa. Haku kohdistuu kaikkiin videoihin, joiden metatiedon sisältämä tilannetieto on lähellä käyttäjän nykyistä tilannetta (kontekstia).

Koodausmenetelmillä joustavuutta

Skaalautuvalla videokoodauksella videoleike jaetaan peruskerrokseen ja yhteen tai useampaan parannuskerrokseen. Peruskerros sisältää tehokkaasti pakatun videoleikkeen, jonka välitys tietoverkossa ei vaadi suurta tiedonsiirtokapasiteettia. Parannuskerros parantaa peruskerroksen videokuvan laatua.

GPRS-verkossa tuttavan multimediapuhelimeen lähetetty leike koostuisi ainoastaan peruskerroksesta. Bluetooth-yhteyden yli omaan tietokoneeseen siirretty leike sisältäisi sekä peruskerroksen että parannuskerroksen, jolloin video soveltuisi paremmin tietokoneen näytölle.

Nykyiset skaalautuvat koodausmenetelmät eivät kykene yhtä hyvään kompressiosuhteeseen kuin uusimmat ei-skaalautuvat videokoodausmenetelmät verrattaessa menetelmien toimintaa ennakkoon määrätyllä tiedonsiirtonopeudella. Koodaustehokkuuden kasvattaminen on tutkimustyön alla ja yksi keino parantaa tehokkuutta on käyttää apuna sisältöanalyysiä, joka ilmaisee käyttäjän kannalta kiinnostavan alueen videokuvassa. Kiinnostava alue, kuten ihmisen kasvot, voidaan koodata tarkemmin ja kuvan tausta vastaavasti heikommin kuin perinteisellä videonkoodauksella. Kiinnostavan alueen tarkemmalla koodauksella videokuvan subjektiivinen laatu paranee.

Järjestelmä voidaan saada mukautumaan automaattisesti käyttäjän ja päätelaitteen mukaan käyttämällä erilaisia profilointitekniikoita. Näiden avulla päätelaite ja palvelinjärjestelmä määrittävät parametrit, joiden perusteella valitaan sopiva esitysmuoto. Sisältöä voidaan muokata esimerkiksi pienentämällä kuvien kokoa tai niiden resoluutiota, joissain tapauksissa kuvat voidaan jättää kokonaan pois. Videomateriaali voidaan purkaa yksittäisiksi kuviksi, jotka kuvaavat videon tärkeimmät tapahtumat.

Sisällön muokkaus voidaan suorittaa etukäteen (staattisesti) valitsemalla eri päätelaitelaiteryhmille omat vaihtoehdot tai dynaamisesti vasta palvelupyynnön saavuttua. Molemmissa tavoissa on etunsa ja haittansa: dynaaminen muokkaus lisää palvelimella tarvittavaa laskentaa, mutta pienentää palvelimen tallennuskapasiteetin tarvetta ja staattinen muokkaus lisää palvelimella tarvittavaa tallennuskapasiteettia, mutta saattaa nopeuttaa toimintaa.

Aiheesta enemmän

CANDELA-projekti: www.extra.research.philips.com/euprojects/candela/index.htm
VTT Elektroniikka: www.vtt.fi/ele
Hantro: www.hantro.com
Solid-tietokannat: www.solidtech.com
MPEG-kotisivu: www.chiariglione.org/mpeg/index.htm

Taustat

Kirjoittajat: Johannes Peltola, VTT, Marko Jurvansuu, VTT, Marko Palola, VTT, Jari Korva, VTT, Sari Järvinen, VTT, Antti Heikkinen, VTT, Anne Sachinopoulou, VTT, Johan Plomp, VTT, Jarkko Nisula, Hantro, Timo Hotti, Solid, Jukka Rantala, Solid.

Yhteyshenkilö:

Johannes.Peltola@vtt.fi

Tutkimus: CANDELA

Yhteistyössä: VTT Elektroniikka, Solid Information Technology Oy, Hantro Products Oy, TEKES

Teknologiaohjelma: NETS, EUREKA/ITEA