

SGN-16006, HARJOITUSTYÖ

Antti Altsten

199716

Antti.altsten@student.tut.fi

LYHENNELMÄ

Laskin tässä harjoitustyössä videon kolmelle perättäiselle kuvalle keskimmäisen kuvan suhteen liikevektorikartan ja neliökeskivirheet matlabilla. Työssä toteutettiin harjoitustyön työohjeessa kuvatut pakolliset osiot yhdestä kolmanteen. Opin työssä, että en osaa ilmeisesti laskea keskineliövirhettä oikein, koska tulokseni piti kertoa mystisellä kertoimella 22,5781, jotta ne vastasivat esimerkkivastausta. Opin myös, että uint8 muodossa ei voida esittää lainkaan negatiivisia arvoja. Muut työvaiheet olivat jollain tapaa tuttuja jo kurssilla aikaisemmin suorittamistani harjoituksista tai muilta kursseilta. Työssä hyödynsin paljon for-looppeja ja kuvia joutui indeksoimaan myös usein ja tarkasti, jotta matriisien mittasuhteet vastasivat toisiaan ja laskemisessa olisi ylipäätään mitään järkeä. Lukuarvojen suuruusvertailua piti myös harjoittaa sopivien arvojen valitsemiseksi monien joukosta.

1. TEORIA

Kaksiulotteista kuvaa esittävässä videossa on tyypillisesti kolme ulottuvuutta – vertikaalinen horisontaalinen ja temporaalinen ulottuvuus. Kaikkien kolmen dimension koodaaminen tismalleen absoluuttisen oikeilla arvoilla vie todella paljon muistia. Tämä aiheuttaa kustannuksia sekä videon tallentamisessa, säilyttämisessä, että siirtämisessä. Ihmisen näkökyky on kuitenkin onneksemme sillä tavoin rajoittunut, että osa tästä informaatiosta voidaan hyvinkin poistaa ilman, että ihminen havaitsee sanottavaa eroa alkuperäiseen videoon nähden. Videota voidaan siis pakata häviöttömän pakkauksensa lisäksi myös häviöllisesti havaitun kuvanlaadun tästä merkittävästi kärsimättä. Tässä työssä toteutettu liikevektorikartta (tai -kenttä) auttaa videon temporaalisessa häviöllisessä pakkauksessa. Kaikkia kuvia ei siis tarvitse tallentaa täydellisinä vaan riittää, että tallennetaan osa kuvista ja kuvien välistä suhdetta kuvaava liikevektorikenttä. Liikevektorikartan avulla mahdollistetaan niin sanottu ennustava kompressio, jossa tallennetaan esimerkiksi keskimmäinen kuva ja liikevektorikenttä joka kuvaa kohteiden liikettä keskimmäisen sitä edeltävän ja seuraavan kuvan välillä. Edeltävä ja jälkimmäisin kuva voidaan tällöin approksimoida keskimmäisen kuvan ja liikevektorikartan ja keskineliövirhekartan avulla. Menetelmässä oletetaan että liikkuvien kohteiden kirkkaus on vakio ajan suhteen ja että tilatasossa lähekkäin olevilla pisteillä on taipumus liikkua samalla nopeudella samaan suuntaan. Oletetaan siis, että kaikissa kolmessa kuvassa ovat näkyvissä samat kohteet, jotka liikkuvat toistensa suhteen, eikä kuvaan ilmesty tai siitä poistu kohteita ja että kohteet eivät myöskään muutu itsessään eri näköisiksi esimerkiksi kuvauspaikalla vaikuttavien valaistusolosuhteiden johdosta. Harjoitustyössä liike esitetään lohkopohjaisena ja liikettä etsitään EBMA-algoritmillä etsimällä kullekin lohkolle vastinlohko joko edellisestä tai seuraavasta kuvasta laskemalla etsintälohkon ja vastinlohkon erotuksen keskineliövirheen paikallinen minimi etsintälohkon lähistöltä (neljä kertaa lohkon kokoa isommalta alueelta). Koska osien lohkoista liikkeessä niiden alkuperäiseen sijaintiin ei jää välttämättä mitään arvoja (ellei jokin toinen lohko ole vastaavasti liikkunut sen paikalle) nämä puuttuvat arvot tyypillisesti extrapoloidaan naapuripikseleistä. Suurin vika tällaisessa lohkopohjaisessa liike-estimoinnissa on tyypillisesti liikkuvien lohkojen reunoille syntyvät lohkorajat. Nämä rajat näkyvät terävinä horisontaalisina tai vertikaalisina reunoina, jotka ihmissilmä valitettavasti erottaa hyvin herkästi. Lohkopohjaista liike-estimointia käytetään kuitenkin sangen laajalti, esimerkiksi laajalle levinneissä MPEG-1 ja MPEG-2 standardin mukaisissa videopakkausmenetelmissä.

2. MENETELMÄT

Toteutin harjoitustyön matlabilla. Kohdassa yksi piti ensin jakaa kuva numero kaksi 12x20 tasakokoiseen 30x32 pikselin lohkokoon. Tämä tapahtui lataamalla kuva ensiksi 'imread' -komennolla, joka tallentaa kuvan 'uint8' -muodossa matlab-muuttujaan. Myöhempiä vaiheita varten muutin jo tässä vaiheessa kuvan 'double' -muotoon helpottamaan laskutoimituksia. Lohkot tallensin 'cell' -muuttujaan luoppaamalla kuvan x- ja y-suunnassa halutun suuruisin pikseliväleihin ja tallentamalla kunkin lohko yhteen 'cell':n soluun. Kuviiin yksi ja kaksi tuli lisäksi kohdassa yksi lisätä nollija kuvien reunoille, jotta etsintä myöhemmissä vaiheissa voitiin aloittaa tasakokoisilla etsintälohkoilla aivan kuvien reunoista jatkaen vastakkaiseen reunaan. Tämän toteutin 'padarray' -komennolla joka lisää halutun määrän nollija kuvan sivuille.

Kohdassa kaksi harjoitustyön tehtävänannossa käskettiin muodostamaan etsintälohkon sijainnin ympärille etsintäalue josta parasta vastinlohkoa etsitään laskemalla pienin keskineliövirhe joko edellisen tai seuraavan kuvan etsintäalueelta. Tämä hoitui yksinkertaisesti luoppaamalla taas pikseleittäin kuvia yksi ja kolme ja vertaamalla etsintälohkoa vastinlohkoihin ja laskemalla keskineliövirhe kuvien kolmesta eri väritasosta. Vertasin laskettuja keskineliövirheitä toisiinsa ja poimin pienimmän talteen. Sen jälkeen siirryttiin seuraavan etsintälohkon kimppuun ja toteutettiin sama uudestaan kunnes kaikille etsintälohkoille oli löydetty pienimmän keskineliövirheen toteuttava vastinlohko. Keskineliövirheen laskeminen tapahtui etsintälohkon ja vastinlohkon alkioittaisella vähennyslaskulla alkioittain toiseen korotettuna ja laskemalla lopuksi alkiot yhteen kaikista kolmesta värikanavasta. Päädyin myös kertomaan kaikki keskineliövirheet luvulla 22,5781, koska tällöin sain esimerkkivastausta vastaavat keskineliövirheet. Luvun löysin kokeellisesti eikä minulla ole mitään teoreettista ymmärtämystä miksi sen pitää olla juuri tuon suuruinen.

Kolmanessa vaiheessa tehtävänantoa käskettiin tallentamaan kohdassa kaksi löydetty keskineliövirheet MSEkarttaan ja siirtymät LKV-liikevektorikarttaan. Näin tein.

3. YHTEENVETO

Liikevektorikentästä näemme, että joissain osissa kuvaa verrattuna edelliseen ja seuraavaan kuvaan voidaan havaita liikettä. Eli, että edellisessä tai jälkimmäisessä kuvassa on nähdään keskimmäistä kuvaa vastaavan kuvanosion liikkuneen x- ja/tai y-akselilla. Tämä tieto helpottaa tiedoston pakkaamista, koska tällöin kaikkia kolmea kuvaa ei tarvitse tallentaa kokonaisuudessaan vaan riittää, että tallennetaan vaikka keskimäinen kuva ja aproksimoidaan edeltävä ja jälkimmäinen kuva käyttäen liikevektorikenttää. Siirretään siis keskimäisen kuvan sisältöä liikevektorikentän osoittamaan suuntaan ja koska kuvia näytetään niin monta sekunnissa ei ihmissilmä kykene erottamaan pieniä virheitä mitä tällainen aproksimaatio saa aikaan. Yksittäisiä kuvia tarkastellessa voimme siis kuitenkin virheet huomata, koska keskineliövirheet osoittavat, ettei siirtymien mukaisella kuvaosien siirtelyllä päästä täysin edeltävää tai jälkimmäistä kuvaa vastaavaan lopputulokseen vaan muodostetut kuvat poikkeavat hieman alkuperäisistä. Työn tekemiseen meni yhteensä tehokasta työaika noin 10 tuntia. Jos yöunien menettämisen aiheuttama yleisen työtehon lasku lasketaan mukaan niin työaika kohoaa ehkä noin pariin vuorokauteen. Ylipäätään olisin toivonut harjoitustyöhön ehkä jonkinäköistä ohjausta, koska nyt sinne jäi esimerkiksi tuo lähdekoodista löytyvä mystinen kerroin 22,5781 jolla päädyin keskineliövirheeni kertomaan, jotta sain esimerkkiratkaisua vastaavat arvot. Miksi ihmeessä jälkimmäinen kuva on liikevektorikentän kolmannessa ulottuvuudessa merkitty kakkosella eikä kolmosella?