Ilkka Ollari

# TEKNIIKOITA DIGITAALISTEN VALOKUVIEN AITOUDEN ARVIOIMISEEN

Tietojärjestelmätieteen
kandidaatintutkielma
28.8.2009

Jyväskylän yliopisto
Tietojenkäsittelytieteiden laitos
Jyväskylä

# TIIVISTELMÄ

Ollari, Ilkka
Tietojärjestelmätieteen kandidaatintutkielma / Ilkka Ollari
Jyväskylä: Jyväskylän yliopisto, 2009, 28 s.

Kandidaatintutkielma

Digitaalisten valokuvien muokkaamisesta on tullut hyvin helppoa ja yleistä, mikä on osaltaan johtanut valokuvien luotettavuuden heikentymiseen. Tässä tutkielmassa käsittelen kirjallisuuskatsauksen pohjalta erilaisia valokuvien aitouden arvioimiseen tarkoitettuja tekniikoita. Tutkin mm. sitä, kuinka luotettavasti kyseisillä tekniikoilla valokuvien aitous voidaan varmistaa, ja minkälaisia rajoitteita eri tekniikoilla on. Tekniikat voidaan jakaa aktiivisiin ja passiivisiin. Ensin mainitusta ryhmästä käsittelen vesileimoja, jotka ovat merkittävin aktiivisten tekniikoiden alalaji. Niiden toiminta perustuu siihen, että kohteena olevaan kuvaan upotetaan jonkinlainen normaalitarkastelussa näkymätön tunnus, jonka avulla kuvan aitous voidaan myöhemmin todentaa. Vesileimatekniikoiden avulla valokuvien aitous voidaan varmistaa suhteellisen luotettavasti ja tarkasti, mutta ne soveltuvat tietenkin ainoastaan vesileimalla varustettuihin kuviin. Näin ollen valtaosa olemassa olevista kuvista jää niiden käyttöalan ulkopuolelle. Vesileimatekniikoista on esitetty lukuisia toteutuksia, mutta esitän tässä tutkielmassa niiden yleisen toimintaperiaatteen sekä tärkeimpiä ominaisuuksia. Passiivisten tekniikoiden ajatuksena puolestaan on se, että tutkittavasta kuvasta ei yleensä tarvita mitään ennakkotietoa, vaan aitouden arviointi perustuu esimerkiksi tilastolliseen päättelyyn. Valokuvien aitous voidaan optimaalisissa olosuhteissa todentaa myös erilaisten passiivisten tekniikoiden avulla, mutta niiden käyttöön liittyy suhteellisen paljon epävarmuustekijöitä. Passiivisia tekniikoita yhdistelemällä voidaan kuitenkin päästä huomattavasti tarkempiin tuloksiin. Esittelen kuusi erilaista passiivista tekniikkaa, jotka voidaan jaotella sen mukaan, perustuvatko ne pikselitason havaintoihin, kameran ominaisuuksiin vai tiettyyn tiedostomuotoon. Sekä vesileimatekniikoita, että passiivisia tekniikoita vastaan on kehitetty erilaisia vastatoimia, eli hyökkäyksiä. Yleisesti ottaen passiiviset tekniikat ovat alttiimpia vastatoimille

AVAINSANAT: Digitaalisten valokuvien autentikointi, kuvan muokkaus, vesileima, digital image forensics

Ohjaaja: Pasi Tyrväinen
Tietojenkäsittelytieteiden laitos
Jyväskylän Yliopisto

Tarkastaja: Timo Käkölä
Tietojenkäsittelytieteiden laitos
Jyväskylän Yliopisto

# SISÄLLYSLUETTELO

# 1 JOHDANTO

Digitaalisten kameroiden määrä on kasvanut räjähdysmäisesti viime vuosien aikana. Myös yhä useammat matkapuhelimet sisältävät nykyään jonkinlaisen kameran. Laadukkaita kuvankäsittelyohjelmia on saatavilla hyvin edullisesti ja jopa täysin maksutta. Ohjelmien käyttö on suhteellisen helppoa ja ohjeita kuvien muokkaamiseen on tarvittaessa saatavilla. Parhaimmillaan muokkaukset ovat niin taidokkaita, ettei niitä välttämättä voi paljaalla silmällä erottaa aidoista kuvista. Muokkauksen helppous ja yleisyys ovat osaltaan heikentäneet luottamusta valokuviin uskottavina todisteina todellisista tapahtumista. Valokuvien merkitys monilla yhteiskunnan alueilla on kuitenkin merkittävä. Valokuvia voidaan käyttää hyväksi muun muassa oikeudellisina todisteina, uutisoinnin apuna tiedotusvälineissä, propagandatarkoituksissa sekä tieteellisten tulosten raportoinnissa. Muokattujen kuvien käytöllä tällaisissa tarkoituksissa voisi olla vakavia seurauksia. Edellä mainitut seikat ovat johtaneet siihen, että tarvitaan tehokkaita keinoja digitaalisten kuvien aitouden selvittämiseen.

Tässä tutkielmassa käsittelen erilaisia tekniikoita, joita alan kirjallisuudessa on ehdotettu digitaalisten valokuvien aitouden arvioimiseen. Tekniikat voidaan jakaa ensinnäkin *aktiivisiin* ja *passiivisiin* (mm. Zhang ym. 2008). Aktiivisista tekniikoista käsittelen vesileimatekniikoita. Vesileimatekniikoiden toiminta perustuu siihen, että kuvaan upotetaan jonkinlainen normaalitarkastelussa näkymätön tunnus, jonka avulla kuvan aitous voidaan myöhemmin todentaa (Cox ym. 1997). Passiivisia tekniikoita ja niiden tutkimusta tarkoittamaan on englanninkielisessä kirjallisuudessa omaksuttu hankalasti suomennettava termi *digital image forensics*. Tällä viitataan ikään kuin digitaalisten kuvien rikospaikkatutkintaan. Tutkimusalue on varsin tuore, mutta se on herättänyt viime vuosina runsaasti mielenkiintoa. Passiivisten tekniikoiden perustavana ajatuksena on se, että tutkittavasta kuvasta ei tarvita pääsääntöisesti mitään ennakkotietoa, vaan aitouden arviointi perustuu esimerkiksi tilastolliseen päättelyyn (Zhang ym. 2008). Vesileimatekniikoita on tutkittu huomattavasti enemmän, mutta niidenkin osalta valtaosa tutkimuksesta on keskittynyt mm. tekijänoikeuksien suojaamiseen tarkoitettuihin vesileimajärjestelmiin (Kundur & Hatzinakos 1999). Kuitenkin viime vuosien aikana sisällön autentikointi on herättänyt kasvavaa mielenkiintoa, ja tähän tarkoitukseen on kehitetty entistä parempia vesileimajärjestelmiä. Uudet toteutukset ovat tarjonneet parannusta paitsi muokattujen alueiden tunnistustarkkuuteen, myös tekniikoiden turvallisuuteen ja luotettavuuteen. Vesileimatekniikoiden oleellisin rajoite on niiden soveltuvuus ainoastaan vesileimalla varustettuihin kuviin. Passiivisilla tekniikoilla ei ole tällaista rajoitetta, mutta niillä saavutettavat tulokset ovat yleensä vesileimatekniikoita epävarmempia. Passiiviset tekniikat ovat myös jonkin verran alttiimpia erilaisille vastatoimille.

Lyhyesti muotoiltuna tutkimusongelma on seuraava: minkälaisia tekniikoita on kehitetty digitaalisten valokuvien aitouden arvioimiseksi, miten hyvin tekniikat soveltuvat tähän tehtävään ja kuinka luotettavasti valokuvien aitous voidaan arvioida kyseisiä tekniikoita apuna käyttäen. Tekniikoiden arvioinnissa käyttämäni kriteerit voidaan jakaa neljään pääkohtaan. 1) *Toimintaperiaate.* Tarkoitukseni ei ole tutkia menetelmien teknisiä yksityiskohtia erityisen tarkalla tasolla, vaan esittää yleiskuva kunkin menetelmän toimintaperiaatteesta ja tyypillisistä ongelmista, joita menetelmien toteutukseen liittyy. 2) *Tunnistuksen kattavuus ja tarkkuus.* Arvioin sitä, minkä tyyppisiä muokkauksia tekniikat pystyvät tunnistamaan ja kuinka tarkasti tunnistus tapahtuu.

Muokkaustyyppien osalta ei tässä yhteydessä ole kuitenkaan mahdollista esittää täysin kattavaa listausta. Tunnistuksen tarkkuuteen liittyy muun muassa se, kuinka tarkasti kullakin menetelmällä voidaan tunnistaa ne alueet, joihin muokkaus kohdistuu. Tunnistustarkkuuteen voidaan laskea myös menetelmän kyky tehdä oikeita tunnistuksia suhteessa virheellisesti tunnistettuihin aitoihin kuviin (false positive) ja tunnistamatta jääneisiin muokattuihin kuviin (false negative). 3) *Rajoitteet.* Pyrin antamaan arvion siitä, kuinka hyvin käsitellyt menetelmät todella soveltuvat erilaisiin käytännön tilanteisiin. Tähän vaikuttavat ainakin menetelmien rajoitukset, kuten sidonnaisuus tiettyyn kuvaformaattiin, suuret laskennalliset vaatimukset, ynnä muut vastaavat seikat. 4) *Luotettavuus.* Menetelmän luotettavuuteen vaikuttaa etenkin se, onko sitä vastaan kehitettävissä tai jo kehitetty vastatekniikoita, eli hyökkäyksiä.

On syytä määritellä mitä valokuvan aitoudella tässä tapauksessa tarkoitetaan, sillä käsite ei ole aivan yksiselitteinen. Voidaan esimerkiksi miettiä onko kuva julistettava välittömästi väärennetyksi, jos se on läpikäynyt mitä tahansa muutoksia. Tämä voisi toki olla yksi tulkinta, mutta käytännössä se lienee hieman liian tiukka kriteeri. Lin ym. (2000) jaottelevat kuvaa muuntavat operaatiot informaation säilyttäviin ja informaatiota muuttaviin. Voidaan siis arvioida onko kuvan sisältämä informaatio tai merkitys säilynyt, vaikka se olisi läpikäynyt tiettyjä muutoksia. Esimerkiksi häviöllinen pakkaaminen on hyvin tavallinen digitaalisille kuville suoritettava toimenpide, joka kylläkin aiheuttaa tiettyjä muutoksia kuvaan, mutta ei ainakaan kohtuullisesti käytettynä muuta kuva sisältämää informaatiota tai merkitystä. Ei ole tietenkään aivan itsestään selvää, missä menee hyväksyttävän ja kielletyn muokkauksen raja. Suurelta osin se lienee päätettävä kuvien käyttötarkoituksen ja kontekstin mukaan. Tietyissä yhteyksissä tavanomaiset kuvan ehostustoimenpiteet, kuten vaikkapa tarkkuuden korostaminen tai värien korjaaminen voivat olla hyväksyttäviä. Joissakin yhteyksissä taas voi olla tärkeää, ettei kuvaan ole kohdistunut minkäänlaisia muutoksia aiheuttavia operaatioita. Sen sijaan tietyt muokkaustoimenpiteet voidaan yleensä suhteellisen yksiselitteisesti luokitella kielletyiksi. Tällaisia ovat ainakin kuvan osien poistaminen tai siirtäminen, uusien osien liittäminen joko toisista kuvista tai ohjelmallisesti generoimalla, sekä muutokset kuvan osien kokoon tai asentoon. Autentikointitekniikoiden näkökulmasta erottelu kuvien hyväksyttävään ja kiellettyyn muokkaamiseen aiheuttaa luonnollisesti lisähaasteen. Kykyä erottaa eri tyyppisiä muokkaustapoja voidaankin pitää tärkeänä arviointikriteerinä tarkastelluille tekniikoille. Toisaalta on mahdollista, että lievempiä prosessointitoimenpiteitä käytetään peittämään varsinaisen muokkauksen aiheuttamia jälkiä. Esimerkiksi häviöllinen pakkaus aiheuttaakin ongelmia etenkin monien passiivisten autentikointimenetelmien toiminnalle.

Luvussa kaksi käsittelen vesileimatekniikoita. Vesileimatekniikoista on kehitetty lukuisia toteutuksia, mutta pyrin käsittelemään aihetta suhteellisen yleisellä tasolla siten, että tekniikoiden oleellisimmat piirteet tulevat esille. Otan kuitenkin esille myös joitakin yksittäisiä toteutuksia, jotka tarjoavat käsitellyn ongelman kannalta erityisen hyödyllisiä ominaisuuksia. Luvussa kolme käsittelen passiivisia tekniikoita. Esittelen yhteensä kuusi erilaista tekniikkaa, joiden käsittelyssä käytän Faridin (2009, 16) esittämää jaottelua. Hänen mukaansa passiiviset tekniikat voidaan jakaa mm. pikselitason havaintoihin, kameran ominaisuuksiin ja tiedostomuotoon perustuviin.

# 2 VESILEIMATEKNIIKAT

Aktiivisten autentikointitekniikoiden osalta keskityn tässä tutkielmassa vesileimatekniikoihin. Niiden lisäksi aktiivisiin tekniikoihin luetaan ainakin digitaaliset allekirjoitukset (Zhang ym. 2008). Jätän digitaaliset allekirjoitukset käsittelyn ulkopuolelle, sillä viimeaikainen tutkimus näyttää keskittyneen voittopuolisesti vesileimoihin. Digitaaliset allekirjoitukset eroavat vesileimoista siinä, että autentikointiin käytettyä dataa ei upoteta kuvainformaatioon, vaan se pidetään erillisessä tiedostossa tai kuvan otsikkotiedoissa (Rey & Dugelay 2002). Tosin kyseisiä käsitteitä käytetään toisinaan myös päällekkäin (Mohanty 1999).

Vesileimalla tarkoitetaan tiettyä dataa, esimerkiksi otsikkoa tai tunnistetta, joka upotetaan kohdeobjektiin siten, että myöhemmin vesileima voidaan tunnistaa ja sen perusteella kohdeobjektista voidaan tehdä erinäisiä päätelmiä. Kohteena voi kuvien lisäksi olla muunkinlaista sisältöä, kuten videota tai audiota. (Mohanty 1999, 4) Eri mediat voivat kuitenkin asettaa käytetylle tekniikalle jossain määrin erilaisia vaatimuksia (Podilchuk & Delp 2001). Yleensä vesileima -järjestelmissä käytetään salausavaimia vesileiman muuttamisen tai poistamisen estämiseksi (Hartung & Kutter 1999). Sen sijaan vesileima, toisin kuin salaus yleensä, ei estä kohteena olevan sisällön tarkastelua (Podilchuk & Delp 2001).

Mohantyn (1999, 5) mukaan vesileimat voidaan jakaa lukuisiin eri kategorioihin niiden eri ominaisuuksien perusteella. Kaikkein oleellisin jakoperuste liittyy vesileimojen käyttötarkoitukseen. *Kestäviä* (robust) vesileimoja on perinteisesti käytetty tekijänoikeuksien suojaamiseen ja sisällön seuraamiseen. Sisällön autentikointiin puolestaan on tavattu käyttää *heikkoja* (fragile) vesileimoja. Kestävien vesileimojen perustavana ajatuksena on se, että niiden tulisi säilyä mahdollisimman hyvin isäntäsignaalissa siihen kohdistuvista muutosoperaatioista huolimatta. Heikkojen vesileimojen toiminta perustuu siihen, että kuvaan kohdistetut muutokset muuttavat myös vesileimaa (Kundur & Hatzinakos 1999).

Kestävien ja heikkojen vesileimojen lisäksi on kehitetty näiden kahden tyypin välimuoto, josta käytetään englanninkielisessä kirjallisuudessa termiä *semi-fragile*. Käytännön sovellusten näkökulmasta heikkojen vesileimojen suurin ongelma on niiden kykenemättömyys erottaa häviöllistä pakkausta ynnä muita ”sallittuja” operaatioita varsinaisesta kuvan muokkaamisesta. (Rey & Dugelay 2002) Muun muassa Lin ym. (2000) ovat esittäneet toteutuksen, joka yhdistelee ominaisuuksia kestävistä ja heikoista vesileimoista. Se kykenee paikantamaan kuvan muokatut alueet, mutta lisäksi se sallii kestävien vesileimojen tapaan jonkin verran mm. häviöllistä pakkaamista.

Samaan objektiin voidaan myös upottaa useita vesileimoja. Tällöin niiden lisäysjärjestyksellä voi olla merkitystä. Tosin ainakin Lu ja Liao (2001) ovat kehittäneet menetelmän, jossa useita vesileimoja voidaan lisätä missä järjestyksessä hyvänsä. Heidän menetelmänsä yhdistää kuvan autentikoinnin ja tekijänoikeuksien suojaamisen.

## 2.1 Vesileimatekniikoiden vaatimukset

Useimmiten vesileimoja käsittelevässä kirjallisuudessa vesileimoille asetetaan kolme perusvaatimusta, jotka ovat havaitsemattomuus, kapasiteetti ja kestävyys (mm.

Wolfgang ym. 1999). Vaatimukset ovat keskenään jossain määrin ristiriitaisia. Vesileiman käyttötarkoitus sanelee osaltaan sen, mitä ominaisuuksia painotetaan. Tässä tarkastelen vaatimuksia erityisesti valokuvien autentikoinnin näkökulmasta. Siinä kontekstissa vesileiman kestävyys ei perinteisesti ole ollut kaikkein tärkein tekijä, mutta se on kuitenkin huomioitava erityisesti sellaisten toteutusten yhteydessä, jotka sallivat pieniä muutoksia, kuten häviöllistä pakkausta, tarkasteltavaan kuvaan.

Kuten todettu, vesileiman kestävyydellä tarkoitetaan vesileiman kykyä säilyä isäntäsignaalissa mahdollisimman hyvin. Ihannetapauksessa vesileima tulisi olla poistettavissa vain sillä ehdolla, että poisto aiheuttaa kuvanlaadulle niin suurta heikkenemistä, että kuvasta tulee käytännöllisesti katsoen käyttökelvoton (Wolfgang ym. 1999). Kapasiteetilla tarkoitetaan sitä, kuinka paljon informaatiota vesileimaan voidaan sisällyttää. Havaitsemattomuus puolestaan tarkoittaa sitä, että vesileiman tulisi olla normaalissa tarkastelussa näkymätön, eikä se saisi heikentää isäntäsignaalin laatua. (Cox & Miller 1997)

Cook ja Rajan (2006) ovat Mohantyn tekstin (1999) pohjalta esittäneet perustavia vaatimuksia erityisesti kuvien autentikointiin tarkoitetuille vesileimajärjestelmille. Järjestelmän tulisi ensinnäkin tunnistaa kaikenlainen kuvan muuttaminen. Lisäksi järjestelmän olisi syytä kyetä osoittamaan suhteellisen tarkasti muutosten sijainti. Vesileima on voitava tunnistaa ilman alkuperäistä kuvaa, sillä sellaista ei välttämättä ole olemassa tai ainakaan yleisesti saatavilla. Tällaista menetelmää kutsutaan julkiseksi vesileimaukseksi tai vesileiman sokeaksi havaitsemiseksi (Kutter & Petitcolas 1999). Vesileima tulisi olla asymmetrisesti, eli julkisen avaimen menetelmällä salattu, jotta se olisi mahdollista tunnistaa ja verifioida julkisesti. Tällä tavoin vesileiman upottamiseen käytettyä avainta ei tarvitse julkistaa. Vesileiman tulee myös olla niin turvallinen, että oikeudettomat osapuolet eivät voi muuttaa tai luoda vesileimaa uudelleen siten, että se vastaisi kuvaan tehtyjä muutoksia. Kerckhoffin periaatteen mukaan salausjärjestelmän turvallisuuden tulee perustua ainoastaan avaimen, ei järjestelmän toteutuksen salaisuuteen. Toisin sanoen vesileiman salaamiseen käytetyn avaimen tulee olla riittävän pitkä.

Edellisten lisäksi Rey ja Dugelay (2002) esittävät lisävaatimuksen, jonka mukaan järjestelmän tulisi ainakin karkealla tasolla kyetä palauttamaan kuvan alkuperäinen informaatio alueelta, jota on muokattu. Käytännössä melko harvat vesileimajärjestelmät toteuttavat tämän vaatimuksen.

## 2.2 Toimintaperiaate

Vesileimajärjestelmän toiminnassa on erotettavissa kolme erillistä vaihetta. Ensimmäinen vaihe käsittää vesileimasignaalin luomisen. Vesileima riippuu yleensä avaimesta ja vesileimainformaatiosta. Vesileimainformaatio voi olla täysin satunnaista, mutta se voi myös riippua esimerkiksi kohteena olevan kuvan sisällöstä. Toisessa vaiheessa vesileima upotetaan isäntäsignaaliin. Toisinaan kahta ensimmäistä vaihetta ei lasketa erillisiksi, etenkin jos vesileiman sisältö riippuu kuvainformaatiosta. Toisen vaiheen tuloksena saadaan vesileimalla varustettu kuva. (Hartung & Kutter 1999) Vesileiman upottamiseen käytetty tapa voi vaihdella huomattavasti eri toteutuksissa. Holliman ja Memon (2000, 432) ovat kuitenkin esittäneet yleisen mallin, jolla

vesileiman upottamista voidaan havainnollistaa. Malli voidaan esittää seuraavalla yhtälöllä:

$$X' = E_K(X, W)$$

Yhtälössä $X$ on alkuperäinen kuva, $X'$ on vesileimattu kuva, $W$ on vesileimainformaatio, $K$ on upottamiseen käytetty avain ja $E$ on upottamiseen käytetty funktio.

Kolmannessa vaiheessa vesileimasignaali tunnistetaan ja puretaan vesileimatusta kuvasta. Tässä yhteydessä kuva voidaan ajatella ei toivottuna signaalina tai kohinana, ja vesileima varsinaisena signaalina. Kuitenkin kuvasignaalin osuus on huomattavasti suurempi kuin vesileiman, joten signaali-kohinasuhde on tällöin paljon pienempi kuin yksi. (Cox & Miller 1997) Hollimanin ja Memonin (2000) mallin mukaan vesileiman purkaminen voidaan esittää seuraavasti:

$$\hat{W} = D_{K'}(\hat{X}')$$

Yhtälössä $\hat{X}'$ on mahdollisesti korruptoitunut, vesileimattu kuva, $K'$ on vesileiman purkamiseen käytetty avain, $D$ on vesileiman purkamiseen käytetty funktio ja $\hat{W}$ on purettu vesileimainformaatio.

Arvio kuvan aitoudesta saadaan vertaamalla alkuperäistä sekä purettua vesileimaa keskenään. Vertailun tarkemmat yksityiskohdat sekä sen perusteella saatava informaatio riippuvat käytettävästä vesileimajärjestelmästä. Useissa järjestelmissä esimerkiksi muokattujen alueiden sijainti on pääteltävissä. (Rey & Dugelay 2002) Joissakin järjestelmissä on myös mahdollista arvioida muutosten astetta (mm. Kundur & Hatzinakos 1999).

Vesileiman upottaminen voi kohdistua joko kuvan normaaliin tila-avaruuteen tai vaihtoehtoisesti johonkin muunnosavaruuteen. Tyypillisiä muunnoksia ovat muun muassa Fourier -muunnos, diskreetti kosinimuunnos ja aallokemuunnos. (Hartung & Kutter 1999) Muunnosavaruuksissa pikseli-informaatio esitetään taajuuksina, jotka edustavat kuvan erityyppisiä alueita (Potdar ym. 2005). Muunnosavaruuksien käyttämisellä on havaittu olevan tiettyjä etuja vesileiman näkymättömyyden ja turvallisuuden suhteen (Hartung & Kutter 1999). Kun kuva muunnetaan takaisin tila-avaruuteen, vesileima levittäytyy epätasaisesti ympäri kuvaa, joten sen luvaton lukeminen tai muokkaaminen vaikeutuu (Hameed ym. 2006). Muunnosavaruuksien käyttö on kuitenkin laskennallisesti vaativampaa (Potdar ym. 2005).

Lähtökohtaisena ongelmana vesileimajärjestelmissä on datan upottaminen kohdesignaaliin siten, että se jättäisi mahdollisimman vähän havaittavia jälkiä. Toisaalta vesileima tulisi olla silti luotettavasti havaittavissa. (Hartung & Kutter 1999) Etenkin heikoissa vesileimajärjestelmissä vesileima on usein upotettu kuvan pikseleiden vähiten merkitseviin bitteihin. Tällä tavoin vesileiman huomaamattomuus on suhteellisen helppo varmistaa. Ongelmana on kuitenkin se, että vähiten merkittävät bitit voivat muuttua esimerkiksi pakkaamisen seurauksena. (Lin ym. 2000) Toisaalta on myöskin täysin mahdollista muuttaa kuvaa siten, että vähiten merkitsevien bittien arvot eivät muutu (Kundur & Hatzinakos 1999). Kestäviin vesileimoihin perustuvissa järjestelmissä vesileima sijoitetaan sen sijaan kuvainformaation merkitseviin bitteihin. Tällä tavoin pyritään estämään vesileiman tahallinen tai tahaton poistaminen. (Cox & Miller 1997) Muunnosavaruuksiin kohdistuvissa järjestelmissä vesileima sijoitetaan

yleensä kuvan keskitaajuuksiin. Korkeimpiin taajuuksiin sijoitetut vesileimat voivat kadota esimerkiksi häviöllisen pakkauksen seurauksena ja matalimpiin taajuuksiin kohdistetut muutokset ovat puolestaan todennäköisesti liian näkyviä. (Podilchuk & Zeng 1998) Jotta kuvainformaation merkitseviin osiin upotettu vesileima olisi mahdollisimman huomaamaton, täytyy kuvaan tehtävien muutosten olla pieniä. Pienet muutokset ovat kuitenkin herkkiä muun muassa kohinalle. Ratkaisumalli ongelmaan on lainattu tietoliikennetekniikassa käytettävistä hajaspektritekniikoista. (Servetto ym. 1998) Käytännössä tämä tarkoittaa sitä, että monet vesileimajärjestelmät sijoittavat yhden informaatiobitin toisteisesti useiden pikseleiden tai transformaatiokerrointen alueelle. (Podilchuk & Delp 2001) Tällä tavoin vesileima säilyy huomaamattomana, mutta se voidaan silti tunnistaa luotettavasti (Hartung & Kutter 1999).

Usein näkymättömyyden varmistamiseksi käytetään lisäksi jotakin ihmisen havaitsemiseen perustuvaa mallia (HVS, Human Visual System). Samoja periaatteita on hyödynnetty myös häviöllisten pakkausmenetelmien kehityksessä. (Wolfgang ym. 1999) Kyseiset mallit käyttävät hyväkseen ihmisten epätäydellistä aistijärjestelmää. Visuaalisten signaalien tulee ylittää tietty intensiteetti, ennen kuin ne ovat ihmissilmän havaittavissa. Optimaalisinta olisi upottaa vesileima juuri havaitsemiskynnyksen alapuolelle, mutta sen määritteleminen todellisissa valokuvissa voi olla hankalaa. (Hartung & Kutter 1999) Tässä yhteydessä käytetään yleensä termiä *just noticeable difference* (JND), eli juuri havaittavissa oleva ero. Visuaaliset mallit auttavat nimenomaan tämän kynnyksen määrittelemisessä. Osa malleista hyödyntää kuvan paikallisia ominaisuuksia, jolloin vesileiman kestävyys ja huomaamattomuus voidaan maksimoida. Käytännössä rakenteeltaan monimutkaisille alueille voidaan sijoittaa enemmän informaatiota kuin tasaisille alueille. (Wolfgang ym. 1999)

## 2.3 Erilaisia toteutustyyppejä

Useat vesileimajärjestelmät perustuvat kuvan lohkomiseen. Tähän malliin perustuvissa järjestelmissä vesileiman upottaminen ja purkaminen tapahtuu lohkotasolla. Tällaiset ratkaisut voivat olla edullisia muun muassa laskennallisen tehokkuuden kannalta. Myös muokkauksen paikantaminen on luontevaa toteuttaa tällä tavoin. (Holliman & Memon 2000) Eräs ensimmäisiä ja tunnetuimpia lohkotason vesileimajärjestelmiä on Wongin (1998) esittelemä toteutus. Sittemmin siinä on havaittu erinäisiä turvallisuusongelmia, mutta se on toiminut pohjana monille myöhemmille vesileimajärjestelmille (mm. Fridrich 2002).

Muutamat vesileimatoteutukset tarjoavat mahdollisuuden palauttaa jollakin tasolla alkuperäinen kuvainformaatio muokatuilta alueilta. Tavallisesti tämä toteutetaan siten, että kuvaan upotetaan matalaresoluutioinen versio itsestään. Jokaisen kuvan alueen informaatio tulee kuitenkin upottaa riittävän kauas alkuperäisestä alueesta, jotta sekä varsinainen kuvainformaatio, että vastaava vesileimainformaatio eivät tuhoudu samanaikaisesti. Molempien tietojen tuhoutuminen on tästäkin huolimatta mahdollista, mikäli useampia kuvan alueita muokataan. (Rey & Dugelay 2002)

Joissakin vesileimajärjestelmissä kuvasta tunnistetaan tiettyjä tunnusomaisia piirteitä tai ominaisuuksia, jotka muodostavat vesileiman sisällön ja toimivat pohjana kun arvioidaan kuvan aitoutta. Piirteet voivat koostua vaikkapa kuvan rajoja, värejä tai

kirkkautta koskevasta informaatiosta. Valitut piirteet voivat riippua esimerkiksi tarkasteltavien kuvien sekä kiinnostuksen kohteena olevien muokkausten tyypistä. Piirteet valitaan yleensä sillä perusteella, että ne säilyvät pienistä kuvaan kohdistuvista muutoksista huolimatta, mutta paljastavat oleelliset muutokset. Tällaisia piirteitä voitaisiin mahdollisesti käyttää myös muokattujen alueiden osittaiseen palauttamiseen. Edellä kuvattuja ratkaisuja rajoittaa yleensä se, kuinka paljon informaatiota vesileimaan voidaan sijoittaa. (Rey & Dugelay 2002)

## 2.4 Tunnistuksen tarkkuus

Ensimmäiset vesileimatekniikat mahdollistivat ainoastaan dikotomisen päättelyn muokkauksen tunnistamisessa; joko kuvaa on muokattu tai sitten ei (Kundur & Hatzinakos 1999). Sittemmin tunnistuksen tarkkuudessa on päästy erittäin hyviin tuloksiin, ja muokatun alueen paikantaminen on jo perusvaatimus vesileimatekniikoissa. Parhaat toteutukset kykenevät osoittamaan muokatun alueen jopa pikselin tarkkuudella (mm. Yong-Zhong He & Zhen Han 2008).

Vesileimatekniikat soveltuvat periaatteessa kaikenlaisten muokkausten havaitsemiseen. Käytetyn muokkaustyypin päätteleminen ei kuitenkaan pääsääntöisesti ole mahdollista. Joissakin toteutuksissa tunnistuksen herkkyyttä on kuitenkin mahdollista säätää siten, että häviöllinen pakkaaminen ja muut vähäiset muutokset voidaan erottaa kuvan varsinaisesta muokkaamisesta (mm. Kundur & Hatzinakos 1999).

## 2.5 Rajoitteet

Vaikka vesileimatekniikoiden avulla kuvien autentikointi voidaan tehdä sangen luotettavasti, liittyy niiden käyttöön kuitenkin tiettyjä rajoituksia. Arvioitaessa vesileimatekniikoiden soveltuvuutta, on huomioitava ajateltu käyttötarkoitus. Eräs rajoittava tekijä voi olla esimerkiksi se, että vesileimat jättävät tavallisesti jonkinlaisen jäljen kuvaan. Vaikka jäljet ovatkin normaalitarkastelussa käytännöllisesti katsoen huomaamattomia, voivat ne estää vesileimojen käytön joissakin erityisissä tarkoituksissa, joissa kuvien tarkkuudella on poikkeuksellisen suuri merkitys. Näin voi olla esimerkiksi sen vuoksi, että kuvia joudutaan suurentamaan huomattavia määriä. Tällainen sovellusalue voi olla esimerkiksi lääketiede. (Fridrich 2002) Erääksi ratkaisuksi ongelmaan on esitetty niin sanottu pyyhkiytyvä vesileima. Ajatuksena on se, että vesileima olisi mahdollista poistaa kuvasta purkamisen yhteydessä ja tuloksena saataisiin tällöin alkuperäinen kuvainformaatio. On kuitenkin huomattava, että tällainen ratkaisu saattaisi tarjota varsin houkuttelevan kanavan mahdolliselle hyökkääjälle. (Rey & Dugelay 2002)

Käytännön näkökulmasta kaikkein ilmeisin ongelma vesileimatekniikoiden soveltuvuudessa valokuvien autentikointiin on se, että vesileiman lisäämisen mahdollisuus markkinoilla olevissa kameroissa on hyvin poikkeuksellista. Jotta valokuvien aitous voitaisiin tehokkaasti taata vesileimatekniikoiden avulla, lienee välttämättömänä edellytyksenä se, että vesileima lisätään kuvaan jo ennen kuin se tallennetaan kameran muistikortille tai vastaavalle tallennusmedialle (Farid 2009). On toki mahdollista käyttää erillistä ohjelmaa vesileiman asettamiseen siinä vaiheessa kun

kuva on jo tallennettu. Tällöin on kuitenkin aina olemassa se mahdollisuus, että kuvaa muokataan ennen vesileiman asettamista.

Mikäli vesileiman asettaminen tapahtuu kameran laitteiston toimesta, on huomioitava mahdollisesti rajallinen laskentateho. Toinen oleellinen havainto on se, että tällaisessa tapauksessa vesileimatonta kuvaa ei välttämättä synny lainkaan, jolloin vesileiman huomaamattomuus muodostuu erityisen tärkeäksi kysymykseksi. (Fridrich 1998)

Vaikka vesileimatekniikat eivät yleistyisikään tavallisiin kuluttajakameroihin, niillä saattaisi silti olla joitakin erityisiä sovelluskohteita. Vesileimatekniikalla varustettuja kameroita voitaisiin mahdollisesti käyttää ainakin tilanteissa, joissa kuvien aitous on erityisen tärkeää. Eräänä sovelluskohteena voisi olla vaikkapa rikostutkinta.

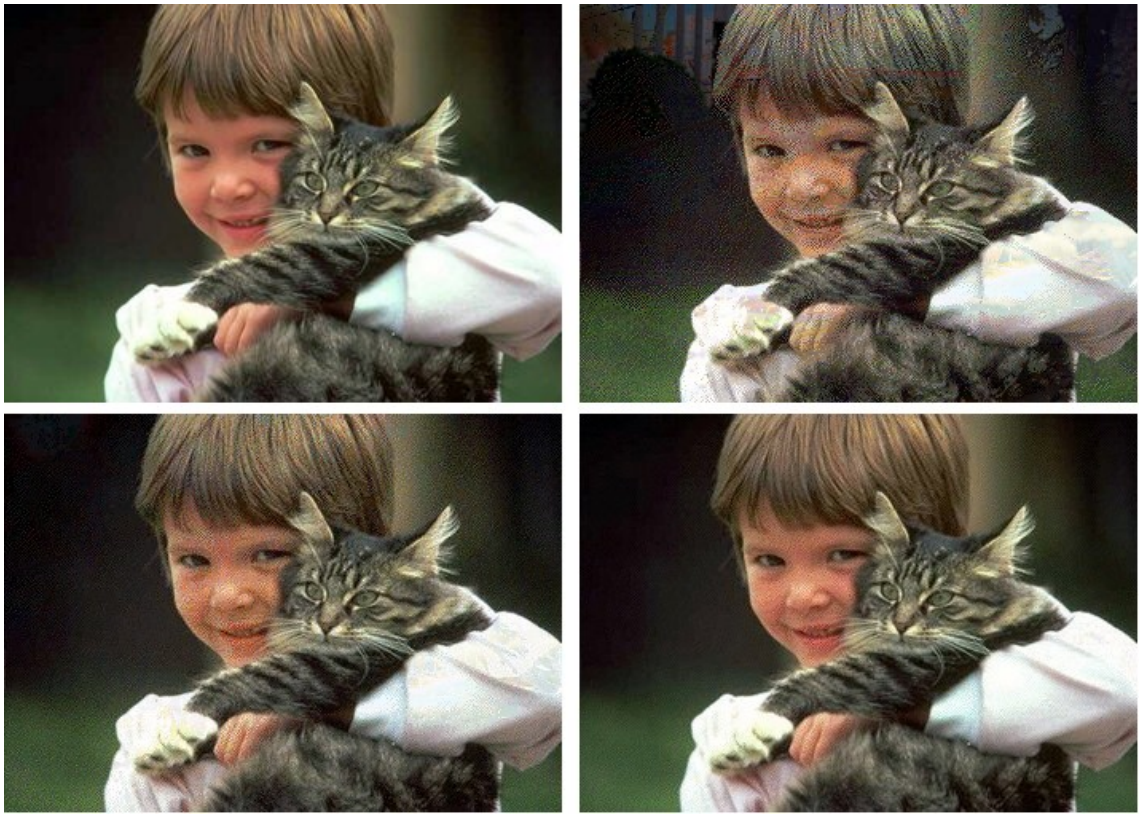
## 2.6 Luotettavuus

Vesileimatekniikoiden luotettavuutta on tutkittu suhteellisen paljon. Monissa vesileimajärjestelmissä on havaittu erinäisiä aukkoja turvallisuuden suhteen, ja entistä parempia, kehittäjiensä mukaan turvallisia toteutuksia on esitetty. (mm. Fridrich 2002, Li 2004) Lienee kuitenkin jossain määrin epäselvää, voidaanko vesileimatekniikoiden turvallisuus yksiselitteisesti taata (mm. Popescu 2005, 8).

Vesileimajärjestelmiä vastaan on kehitetty lukuisia erilaisia hyökkäystapoja. Hyökkäyksen tarkoitus ja toimintaperiaate riippuvat siitä, minkä tyyppiseen vesileimajärjestelmään ne kohdistuvat. Kestäviä vesileimoja vastaan suunnattujen hyökkäysten tarkoituksena on yleensä poistaa vesileima aiheuttamatta merkittävää kuvanlaadun heikkenemistä. Heikkojen ja muiden sisällön autentikointiin tarkoitettujen vesileimojen tapauksessa vesileimaa ei haluta poistaa. Sen sijaan hyökkääjä pyrkii muuttamaan kuvaa ilman, että vesileima muuttuu. Vaihtoehtoisesti tarkoituksena voi olla uuden, aidolta vaikuttavan vesileiman luominen. (Rey & Dugelay 2002)

Vesileiman upottamiseen käytetty avain on tietenkin mahdollista yrittää selvittää raa'alla voimalla. Jos avain saadaan selville, on vesileiman väärentäminen suhteellisen helppoa. Ainoa tapa varautua tällaista hyökkäystä vastaan on käyttää riittävän pitkää avainta, jotta raa'an voiman käyttäminen ei olisi mielekästä. (Rey & Dugelay 2002)

Varsin yleinen heikkoja vesileimoja vastaan käytetty hyökkäys on ns. *vektorikvantisaatio-hyökkäys* (VQ attack, birthday attack, collage attack). Hyökkäyksen kohteena ovat vesileimajärjestelmät, jotka perustuvat kuvan lohkomiseen, mutta eivät edellytä minkäänlaista riippuvuutta lohkojen välillä. Hyökkäyksen periaatteena on koostaa uusi vesileimattu kuva useista kuvista, joissa on käytetty samaa avainta ja vesileimaa, siten, että kuvan osien suhteellinen sijainti säilyy samana. (Li 2004) Hyökkääjän ei tarvitse tuntea vesileiman upottamiseen käytettyä avainta eikä vesileiman sisältöä (Fridrich ym. 2000). Kyseisen hyökkäyksen toteutuksessa suurin hankaluus on saada riittävän iso kokoelma samalla vesileimalla ja avaimella merkittyjä kuvia (Rey & Dugelay 2002). Tosin Fridrichin ym. (2000) testien perusteella jopa 20 kuvan kokoelma riittää väärennetyllä vesileimalla varustetun kuvan koostamiseen. Kuvan laatu on toki tällöin alkuperäistä heikompi, mutta väärennösten laatu paranee kuvien määrän lisääntyessä (Kuva 1). Yleisenä keinona vektorikvantisaatio-hyökkäyksen torjuntaan on poistaa jollakin tavalla lohkojen keskinäinen riippumattomuus (Li 2004). Myös muita keinoja hyökkäyksen torjuntaan on esitetty (mm. Fridrich 2002).



*Kuva 1: Vektorikvantisaatio-menetelmällä koostettuja kuvia. Vasemmalla ylhäällä on alkuperäinen vesileimaton kuva. Muut kuvat on muodostettu 20:n (oikealla ylhäällä), 80:n (vasemmalla alhaalla) ja 300:n (oikealla alhaalla) kuvan kokoelmien avulla. (Fridrich ym. 2000, 6-8)*

# 3 PASSIIVISET TEKNIIKAT

Viime vuosina digitaalisten valokuvien autentikointiin on kehitetty monenlaisia passiivisia tekniikoita, jotka eivät pääsääntöisesti edellytä minkäänlaista ennakkotietoa tutkittavista kuvista. Tällaisten tekniikoiden avulla päästään parhaimmillaan varsin hyviin tuloksiin, mutta kaikilla niistä on kuitenkin omat rajoituksensa. Tässä yhteydessä kattavan esityksen tekeminen kaikista tekniikoista ei ole mahdollista. Olen pyrkinyt valitsemaan tähän tarkasteluun kaikkein varteenotettavimpia tekniikoita. Toisaalta olen valinnut tekniikoita, jotka soveltuvat jossain määrin eri tyyppisiin tarkoituksiin ja näin ollen voivat mahdollisesti kompensoida toistensa puutteita.

Arvioitaessa kunkin käsitellyn tekniikan tarkkuutta, voidaan erottaa kaksi puolta. Ensinnäkin voidaan arvioida menetelmän kykyä tehdä oikeita päätöksiä sen suhteen, onko tarkasteltavaa kuvaa muokattu vai ei. Toisaalta voidaan arvioida sitä, kuinka tarkasti eri tekniikat kykenevät rajaamaan kuvasta ne alueet joihin muokkaus kohdistuu. Tällaisten arvioiden tekeminen ei kuitenkaan ole aivan yksinkertaista, sillä asiaan vaikuttavat hyvin monet tekijät, kuten käytetyn muokkauksen tyyppi ja aste, sekä tietenkin käytetyn autentikointitekniikan erityispiirteet. Tekniikoita ei ole myöskään testattu täysin yhteneväisin periaattein, vaan kehittäjien suorittamat testit poikkeavat luonnollisesti jonkin verran toisistaan. Useimmissa tapauksissa menetelmiä on kuitenkin testattu häviöllisesti pakattujen kuvien suhteen useilla eri pakkauslaaduilla. Kaikissa tapauksissa käytetty tiedostomuoto on tällöin JPEG (Joint Photographic Experts Group), jossa pakkauslaatu on valittavissa asteikolta 1-100. Joissain tapauksissa myös JPEG 2000 -muotoa on käytetty. JPEG 2000 on alkuperäisen JPEG -standardin kehittäjien suunnittelema pakkausstandardi, joka perustuu diskreetin kosinimuunnoksen sijaan aallokemuunnokseen (Skodras, ym. 2001). Pyrin mahdollisuuksien mukaan esittämään jonkinlaisia numeerisia arvioita eri tekniikoiden tarkkuudesta, mikäli tällaisia tietoja on saatavilla. Testien tarkemmat yksityiskohdat ja tulokset löytyvät alkuperäislähteistä.

Menetelmien tarkkuuteen vaikuttavat myös tietyt tekniset ratkaisut. Usein muokkausjälkien esiintymistä tutkitaan jakamalla käsiteltävä kuva kiinteän kokoisiin lohkoihin. Tavallisesti lohkon koon merkitys on se, että pienestä lohkosta muokkausjälkien tunnistaminen voi olla hyvin epävarmaa ja virhealtista. Suuremmasta lohkosta tunnistus voidaan tehdä luotettavammin, mutta alueellinen tunnistustarkkuus puolestaan kärsii. Lohkon koko ei kuitenkaan suoraan kerro tunnistuksen alueellista tarkkuutta, sillä lohkot voivat yleensä olla myös osittain päällekkäisiä. Muokatun alueen rajausta voidaan näin tarkentaa useammista lohkoista saatavien tulosten perusteella. Tämä tietenkin lisää laskennallisia kustannuksia. (esim. Chen ym. 2007)

Tässä luvussa ensimmäisenä esitellyt pikselitason tekniikat rajoittuvat muutamien nimenomaisten muokkaustyyppien tunnistamiseen. Loput tekniikat puolestaan ovat siinä mielessä yleisiä, että ne kykenevät ainakin periaatteessa tunnistamaan lähes kaikki mahdolliset kuvaan kohdistuvat muutokset. Kuitenkaan aivan kaikkia muutoksia ei varmastikaan ole mahdollista havaita näilläkään tekniikoilla. Esimerkiksi kuvan rajaus tai värien säätäminen eivät välttämättä aiheuta kuvaan sellaisia tilastollisia muutoksia, jollaisten havaitsemiseen käsitellyt menetelmät yleensä perustuvat. Tekniikoiden kehittäjät eivät kuitenkaan ole kovinkaan kattavasti käsitelleet sitä, mitkä muokkaustyypit ovat todella tunnistettavissa, ja mitkä eivät. Myöskään tässä yhteydessä

en pyri esittämään kattavaa listausta, johtuen muun muassa muokkaustyyppien suuresta lukumäärästä. Voidaan kuitenkin todeta, että kaikkein oleellisimmin kuvan merkitystä muuttavat toimenpiteet, kuten kuva-alueiden korvaaminen, ovat yleensä tunnistettavissa.

## 3.1 Pikselitason tekniikat

Pikselitason autentikointitekniikat tarkastelevat nimensä mukaisesti erilaisia pikselitasolla esiintyviä ilmiöitä. Niiden toiminta ei perustu esimerkiksi olettamuksiin kameran teknisistä ominaisuuksista, vaan enemmänkin erilaisten muokkaustoimenpiteiden aiheuttamien jälkien tunnistamiseen.

### 3.1.1 Uudelleen näytteistämisen tunnistaminen

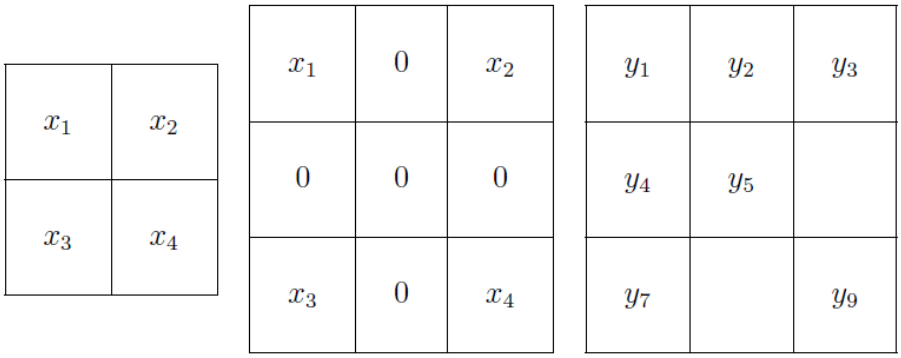
Varsin usein kuvaa muokattaessa joudutaan jotakin kuvan aluetta kiertämään tai venyttämään, taikka sen kokoa joudutaan muuttamaan. Esimerkiksi kuvaan lisättävän henkilön mittasuhteet täytyy mukauttaa kuvan muuhun sisältöön, kuten muihin henkilöihin. Tällaiset operaatiot edellyttävät kuva-alueen *uudelleen näytteistämistä* (resampling), mikä puolestaan yleensä aiheuttaa kuvaan tiettyä jaksoittaista riippuvuutta vierekkäisten pikseleiden välillä. Tällaista riippuvuutta ei yleensä esiinny luonnollisesti, joten sitä voidaan käyttää hyväksi tunnistettaessa jälkiä mm. edellä mainituista muokkaustoimenpiteistä. Esimerkkinä voidaan ajatella yksiulotteista signaalia $x$, joka skaalataan ylöspäin kertoimella kaksi, jolloin saadaan uusi signaali $y$. Tällöin $y$:n parittomat näytteet saavat alkuperäiset $x$:n arvot ja parillisten näytteiden arvoksi muodostuu niiden viereisten näytteiden keskiarvo. Niinpä uudelleen näytteistetty signaali voidaan tunnistaa toteamalla, että sen joka toinen näyte korreloi täydellisesti naapuriarvojensa kanssa. Sama ajatus laajenee hyvin yksinkertaisesti kaksiulotteiseen kuva-avaruuteen. (Farid 2009)

Popescun (2005, 31-37) suorittamien testien perusteella menetelmä on suhteellisen vakaa kuvaan lisättyä kohinaa vastaan. Sen sijaan etenkin JPEG -pakkaus näyttäisi olevan menetelmän heikkous. Tunnistustarkkuus on hyvä silloin, kun pakkauksen laatu on 97:n ja 100:n välillä. Kuitenkin laadun ollessa 90, tunnistustarkkuus putoaa jo lähes sattuman tasolle. Yleisesti ottaen menetelmän tunnistustarkkuuteen vaikuttaa uudelleen näytteistämisen aste.

Menetelmään liittyy myös muita rajoituksia. Ensinnäkin tietyt uudelleen näytteistämisen asteet eivät aiheuta lainkaan jaksoittaista riippuvuutta pikseleiden välille. Toinen menetelmään liittyvä rajoitus on tietenkin se, että sen avulla ei voida tunnista sellaisia muokkaustoimenpiteitä, joihin ei liity uudelleen näytteistämistä. Toisaalta samasta syystä havaitun muokkauksen tyyppi on mahdollista rajata, ja myös uudelleen näytteistämisen aste on tietyin rajoituksin pääteltävissä. (Popescu 2005, 16)

Gloe ym. (2007) ovat esittäneet vastatekniikan, jolla uudelleen näytteistämisen jäljet voidaan peittää varsin tehokkaasti. Menetelmä perustuu karkeasti ottaen siihen, että uudelleen näytteistämisen yhteydessä kunkin pikselin sijaintiin kohdistetaan lievää satunnaista vaihtelua. Periaatteellisella tasolla ajatus on hyvin yksinkertainen, mutta tällaiset toimenpiteet aiheuttavat kuvaan helposti värinää. Oleellista Gloen ym.

tekniikassa on se, että se pyrkii jättämään mahdollisimman vähän havaittavia jälkiä sopeuttamalla muutokset kuvan paikallisiin piirteisiin.




*Piirros 1: Kuvassa vasemmalla on kaksiulotteinen pikselikehikko. Keskellä sama kehikko on skaalattu ylöspäin kertoimella kaksi, jolloin arvot nollalla merkityissä kohdissa jäävät puuttumaan. Oikealla kehikko on osittain uudelleennäytteistettynä. (Popescu 2005, 18)*

### 3.1.2 Monistettujen alueiden tunnistaminen

Eräs tyypillisimpiä muokkaustapoja on peittää jokin kuvan alue, esimerkiksi henkilö tai esine, siten, että peitettävän alueen päälle kopioidaan jokin toinen alue muualta samasta kuvasta. Tällaisia alueita voidaan periaatteessa tunnistaa automaattisesti, mutta ongelmana on se, että muokatun alueen koko, muoto ja sijainti ovat tuntemattomia. Kaikkien mahdollisten vaihtoehtojen tutkiminen olisi kuitenkin laskennallisesti mahdotonta. (Farid 2009) Toinen ongelma kyseisessä lähestymistavassa on se, että monistettua alue ei välttämättä ole täysin identtinen alkuperäisen alueen kanssa. Tällöin on ratkaistava ovatko kaksi aluetta riittävässä määrin samanlaisia ja toisaalta millä perusteella kyseinen päätös tehdään. (Zhang 2008)

Popescu (2005, 91-102) on esittänyt toteutuksen, jossa kuva jaetaan ensin kiinteän kokoisiin lohkoihin, jonka jälkeen lohkoista erotetaan keskeisimmät piirteet niin sanotulla pääkomponenttianalyysillä. Pääkomponenttianalyysi on muun muassa hahmontunnistuksessakin käytetty tilastollinen menetelmä, jonka avulla suuri joukko muuttujia pyritään korvaamaan pienemmällä joukolla uusia muuttujia, jotka säilyttävät mahdollisimman paljon alkuperäisten muuttujien informaatiosta (Yunfei & Ping 2007). Tällä tavoin epäoleelliset, esimerkiksi kohinan tai pakkauksen aiheuttamat eroavaisuudet voidaan sivuuttaa. Lopulta lohkot järjestetään piirteidensä perusteella siten, että samankaltaiset lohkot päätyvät vierekkäin. Lohkojen järjestäminen on laskennalliselta kannalta koko menetelmän vaativin vaihe. Koko algoritmi on aikavaativuudeltaan luokkaa $O(N_t N \log N)$, jossa $N_t$ on pääkomponenttianalyysillä redusoitujen alueiden dimensio ja N kuvan pikseleiden lukumäärä.

Popescun (2005, 93-100) testien mukaan menetelmän tunnistustarkkuus JPEG -muotoisten kuvien suhteen on hyvä jopa pakkauslaadulla 50, silloin, kun tunnistuksessa käytetään suurehkoa lohkokokoa (160 x 160 pikseliä). Hyvin pieniä (32 x 32 pikseliä) lohkoja käytettäessä pakkauslaadun tulee sen sijaan olla lähes täydellinen. Tulokset eivät merkittävästi poikkea JPEG 2000 -muotoisten kuvien osalta, joskin pieniä eroavaisuuksia on. Menetelmä on myös sangen vakaa kuvaan lisättyä kohinaa vastaan. Pienimpiä lohkokokoja lukuun ottamatta tunnistustarkkuus on lähellä täydellistä. Kaikissa tilanteissa väärien hälytysten määrä on melko alhainen, korkeimmillaankin vain noin 3 prosenttia.

Kyseisen menetelmän käyttöala rajoittuu luonnollisesti tilanteisiin, joissa alueita on kopioitu yhden kuvan sisällä paikasta toiseen.

## 3.2 Kameran ominaisuuksiin perustuvat tekniikat

Kameran optiikka sekä eri vaiheissa kuvan prosessointiin käytetyt komponentit jättävät valokuvaan oman jälkensä, jota voidaan käyttää paitsi muokkausjälkien havaitsemiseen, myös tietyn kuvan ottamiseen käytetyn kameran yksilöimiseen.

Digitaalisen kuvan tallentamisprosessi on melko monimutkainen ja vaihtelee huomattavasti eri kameramallien välillä. Kuitenkin tietyt komponentit ja prosessointivaiheet ovat yhteisiä lähes kaikille digitaalisille kameroille. Useimmissa kameroissa on yksittäinen *CCD* (charge-coupled device), tai *CMOS* (complementary metal oxide semiconductor) -kenno. Kenno koostuu pikseleistä, jotka tallentavat valoa muuntamalla fotoneita elektroneiksi. Näin kertynyt sähköinen varaus vahvistetaan ja muunnetaan sitten digitaaliseksi signaaliksi. Digitaalinen värikuva koostuu kolmesta kanavasta, jotka sisältävät näytteitä värispektrin eri taajuusalueilta. Useimmissa kennoissa pikselit eivät kuitenkaan rekisteröi väriä, joten ne on varustettu värisuotimella, joka päästää läpi ainoastaan yhtä väriä; punaista, vihreää tai sinistä. Näin ollen vain yksi näistä väreistä tallennetaan yhden pikselin alueelta. Täydellisen värikuvan muodostamiseksi kaksi muuta väriä on arvioitava viereisten näytteiden perusteella. Tästä toimenpiteestä käytetään nimeä *interpolaatio*. Lopulta kuvaan voidaan kohdistaa erinäistä prosessointia, kuten värien korjailua ja mahdollisesti kohinan poistoa tai tarkennusta. Usein kuva tallennetaan käyttäen jotakin häviöllistä tiedostomuotoa, joista yleisin on JPEG. (Fridrich 2009)

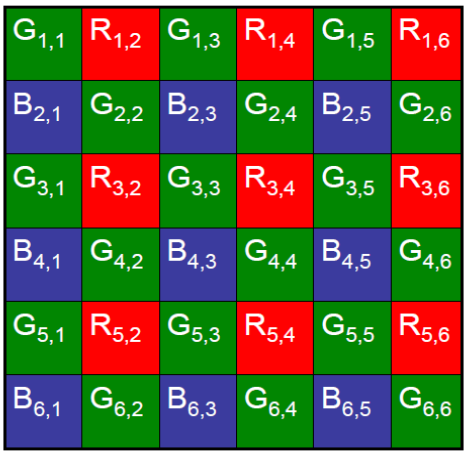
Kameraperustaiset autentikointitekniikat ovat luonnollisesti jossain määrin sidoksissa olemassa olevaan kamerateknologiaan. Sen vuoksi on vaikeaa arvioida, miten kameralaitteistojen kehitys vaikuttaa kyseisten autentikointitekniikoiden käytettävyyteen. Esimerkiksi seuraavana käsiteltävä interpolaatioon perustuva menetelmä ei sovellu aivan kaikille kameroille.

### 3.2.1 Värisuotimen interpolaatio (CFA interpolation)

Popescu ja Farid (2005) ovat esitelleet menetelmän, jossa interpolaatiota voidaan käyttää hyväksi kuvan muokkausjälkien havaitsemisessa. Tavallisimmin värisuotimissa käytetään niin sanottua Bayer-matriisia (Piirros 2), jossa vihreästä valosta tallennetaan 50 prosenttia, sinisestä ja punaisesta puolestaan 25 prosenttia kustakin. Koska puuttuvat

näytteet joudutaan arvioimaan naapuriarvojen perusteella, aiheuttaa interpolointi tietynlaista jaksoittaista korrelointia pikseleiden välille. Kuvaa muokattaessa tällaiset korrelaatiot hyvin todennäköisesti katoavat tai muuttuvat. Interpolointiin käytettäviä algoritmeja on kuitenkin monenlaisia, toiset yksinkertaisempia, toiset monimutkaisempia. Menetelmän haasteena on se, että lähtökohtaisesti ei ole tiedossa, minkälaista algoritmia interpolointiin on käytetty ja mitkä pikselit korreloivat keskenään.

Pakkaamattomien kuvien tunnistamisessa interpolaatiomenetelmä on Popescun (2005, 66-67) testien mukaan erittäin tarkka. Interpolaatioalgoritmista riippuen tunnistustarkkuus on jopa 97-100 prosenttia. Testauksessa käytetyillä asetuksilla vääriä hälytyksiä ei esiintynyt lainkaan. Myöskään kevyt JPEG -pakkaus ei ole tunnistustarkkuuden kannalta kohtalokasta, vaikkakin interpolaatioon käytetty algoritmi vaikuttaa jonkin verran tähänkin seikkaan. Esimerkiksi pakkauslaadulla 70 menetelmän tunnistustarkkuuden vaihtelu on jo huomattavaa (6-56 prosenttia riippuen interpolaatioalgoritmista). Sen sijaan eräs menetelmän suurimmista heikkouksista on se, että se ei sovellu lainkaan JPEG 2000 -menetelmällä pakattujen kuvien autentikointiin. Kyseinen pakkausmetodi muodostaa kuvaan jälkiä, joita ei voi erottaa interpolaatiosta.




*Piirros 2: Bayer-matriisin rakenne. Useimpien digitaalisten kameroiden värisuotimissa käytetään bayer-matriisia. (Gallagher & Chen 2008, 1)*

Menetelmä ei myöskään luonnollisesti sovellu sellaisilla kameroilla otettuihin kuviin, joissa ei ole värisuodinta. Useimmissa digitaalikameroissa sellainen on, mutta poikkeuksen muodostaa ainakin Sigma Corporationin kehittämä Foveon x3 -kenno, joka tallentaa kolmikerroksisen rakenteensa ansiosta kaikki kolme päväriä jokaisen pikselin alueelta (Popescu 2005, 67).

Mahdollinen hyökkäystapa menetelmää vastaan voisi olla muokatun alueen interpolointi ohjelmallisesti. Tällöin hyökkääjän tulee kuitenkin tuntea kuvan interpolointiin alunperin käytetty algoritmi. (Popescu 2005, 68)

### 3.2.2 Kohina

Kaikissa digitaalisissa valokuvissa on tietty määrä kohinaa, jota aiheuttavat erinäiset kameran komponentit ja kuvan prosessointivaiheet. Osa kohinasta on täysin satunnaista ja vaihtelee otoksittain, osa puolestaan on kamerakohtaista ja otoksesta toiseen pysyvää. Valokuvien autentikoinnissa kohinaa voidaan hyödyntää eri tavoin. (Fridrich 2009).

Fridrichin (2009, 26-27) mukaan oleellisin osa kohinasta johtuu kuvakennosta. Hän käyttää siitä termiä *photo-response nonuniformity* (PRNU). Ilmiö johtuu kennon pikseleiden lievästi erilaisesta kyvystä muuttaa fotoneita elektroneiksi. Tämä puolestaan johtuu epätäydellisestä valmistusprosessista sekä kennon valmistusmateriaalina käytetyn piin luonteenomaisesta epätasalaatuisuudesta. Jokaiseen kuvaan muodostuu kennon jättämä heikko kohinakuvio. Kuvio voidaan ajatella eräänlaisena sormenjälkenä, jonka perusteella kuvan aitoudesta voidaan tehdä päätelmiä, sekä lisäksi tietyn kuvan ottamiseen käytetty kamera voidaan yksilöidä. Fridrichin mukaan tällaisella sormenjäljellä on monia hyödyllisiä ominaisuuksia. Se on ensinnäkin satunnaisesta luonteestaan johtuen yksilöllinen jokaiselle kennolle. Sormenjälki myös esiintyy kaikissa kennoissa sekä kaikissa kuvissa, riippumatta kameran optiikasta tai asetuksista, eikä se juurikaan muutu ajan tai olosuhteiden vaikutuksesta. Lisäksi sormenjälki kestää kohtuullisen hyvin erilaisia prosessointitoimenpiteitä, kuten pakkausta.

PRNU -jälkeä voidaan ajatella ikään kuin kameran tahattomasti kuvaan jättämänä vesileimana (Chen ym. 2007). Erona on tietenkin se, että tässä tapauksessa ”vesileiman” sisältö on täysin satunnainen, eikä sen tarkkaa sijaintia tunneta. Jäljen erottamiseksi kuvasta muodostetaan sopivan suodattimen avulla kohinaton versio. Arvio kohinasta saadaan tällöin vertaamalla alkuperäistä ja kohinatonta kuvaa (Chen ym. 2007). Koska kaikissa kuvissa on myös täysin satunnaista kohinaa, ei PRNU -komponentin erottaminen onnistu yksittäisestä kuvasta. Hyvä arvio sormenjäljestä vaatii 20-50 muokkaamatonta kuvaa (Fridrich 2009).

Chen ym. (2007) ovat esittäneet PRNU -jälkeen perustuvan menetelmän kuvien autentikointiin. Muokattujen alueiden tunnistaminen menetelmällä perustuu siihen, että PRNU -jäljen esiintymistä tutkitaan kuvan eri alueilla. Jos jälki puuttuu tietyltä alueelta, voidaan alue päätellä muokatuksi. Menetelmällä ei voida kuitenkaan havaita sellaisia muokkaustoimenpiteitä, jotka eivät poista PRNU -jälkeä. Tällaisia ovat esimerkiksi kirkkauden, värien tai kontrastin säätäminen.

Chen ym. (2007) ovat testanneet menetelmänsä tarkkuutta JPEG -pakattujen kuvien suhteen. Pakkauksen laadun ollessa 90, menetelmä tunnisti 85 prosentissa tapauksista vähintään kaksi kolmasosaa muokatuista alueista. Laadulla 75 vastaava luku oli 73 prosenttia. Menetelmän tunnistusherkkyyteen voidaan vaikuttaa eräänlaisella kynnysarvolla. Maltillisemmilla arvoilla muokatut alueet voivat jäädä tunnistamatta ja herkemmillä arvoilla aiheettomia tunnistuksia voi esiintyä. Tekijöiden mukaan lähes kaikki virheet tunnistuksessa liittyivät kuitenkin tilanteisiin, joissa kuva sisälsi hyvin tummia ja yleensä rakenteeltaan monimutkaisia alueita. PRNU -jälki ei tavallisesti esiinny tällaisilla alueilla. Tunnistustarkkuuteen vaikuttaa myös käytetty lohkokoko. Pienistä lohkoista on vaikea arvioida luotettavasti PRNU -jäljen esiintymistä, ja suuret lohkot puolestaan heikentävät tunnistuksen alueellista tarkkuutta. Hyvä tasapaino saadaan tekijöiden mukaan 64 x 64 – 128 x 128 kokoisilla lohkoilla.

Useiden muokkaamattomien kuvien tarve heikentää menetelmän yleiskäyttöisyyttä huomattavasti, ja rajoittaa sen tilanteisiin, joissa joko kuvan ottamiseen käytetty kamera tai muita kyseisellä kameralla otettuja kuvia on saatavilla. Tekniikka ei ole myöskään erityisen nopea, sillä yksittäisen kuvan tarkastaminen voi kestää joitakin minuutteja. (Chen ym. 2007) Tekijät eivät ole juurikaan käsitelleet mahdollisia menetelmäänsä kohdistuvia vastatekniikoita. Ei kuitenkaan vaikuttaisi epärealistiselta olettaa, että PRNU -jälki olisi kohtuudella väärennettävissä muokattuun kuvaan, mikäli myös hyökkääjällä on käytössään joukko samalla kameralla otettuja aitoja kuvia.

*Kuva 2: PRNU-menetelmällä tunnistettu muokkaus. Muokkaamaton valokuva on vasemmalla ylhäällä ja sen muokattu versio oikealla ylhäällä. Alhaalla vasemmalla muokattu alue on tunnistettu pakkaamattomasta TIFF-kuvasta ja alhaalla oikealla JPEG-kuvasta, jonka laatu on 75. (Chen ym. 2007, 352)*

Mahdian ja Saic (2009) ovat esittäneet kohinaan perustuvan menetelmän, joka ei vaadi ennakkotietoa tutkittavasta kuvasta tai sen ottamiseen käytetystä kamerasta. Tekniikka perustuu siihen havaintoon, että kohinataso on yleensä tasainen kaikilla kuvan alueilla. Niinpä paikallisia muutoksia kohinan tasossa voidaan käyttää hyväksi muokattujen alueiden tunnistamisessa. Menetelmä jakaa kuvan kohinatasoltaan samankaltaisiin alueisiin. Kohinan estimointiin käytetään aalloke -muunnosta, jolla kuvan korkeat taajuudet voidaan erottaa.

Mahdianin ja Saicin mukaan heidän menetelmänsä on hyödyllinen etenkin tilanteissa, joissa muokkausjälkiä pyritään peittämään lisäämällä kuvaan paikallisesti kohinaa. Myös toisesta kuvasta lisätyt ja ohjelmallisesti generoidut alueet voivat muuttaa kohinan tasoa riittävästi. Sen sijaan on epäselvää, kykeneekö kyseinen tekniikka erottamaan alueita, jotka on kopioitu samasta kuvasta, tai toisesta samalla kameralla otetusta kuvasta. Toisaalta ei liene mahdotonta lisätä muokattuun kuvaan kohinaa siten, että se jakautuu tasaisesti koko kuvan alueelle. Mahdianin ja Saicin mukaan heidän

menetelmänsä suurin ongelma on se, että myös täysin aidot kuvat voivat sisältää alueita, joissa kohinan taso poikkeaa toisistaan. Tällaiset tapaukset voivat aiheuttaa vääriä hälytyksiä.

Mahdianin ja Saicin (2009) testien perusteella heidän menetelmänsä kykenee erottamaan melko tarkasti muutoksia kohinan tasossa, jos tarkasteltavissa kuvissa on käytetty maltillista pakkaustasoa (esim. JPEG pakkauksen taso 90). Tarkkuuteen vaikuttaa kohinan vaihtelun aste sekä se, kuinka suuri on tarkasteltava alue. Jos kohinan vaihtelu on huomattavaa ja tarkasteltava alue on suurehko (128 x 128 pikseliä), ei voimakkaampikaan pakkaus välttämättä haittaa.

### 3.3 Tiedostomuotoon perustuvat tekniikat

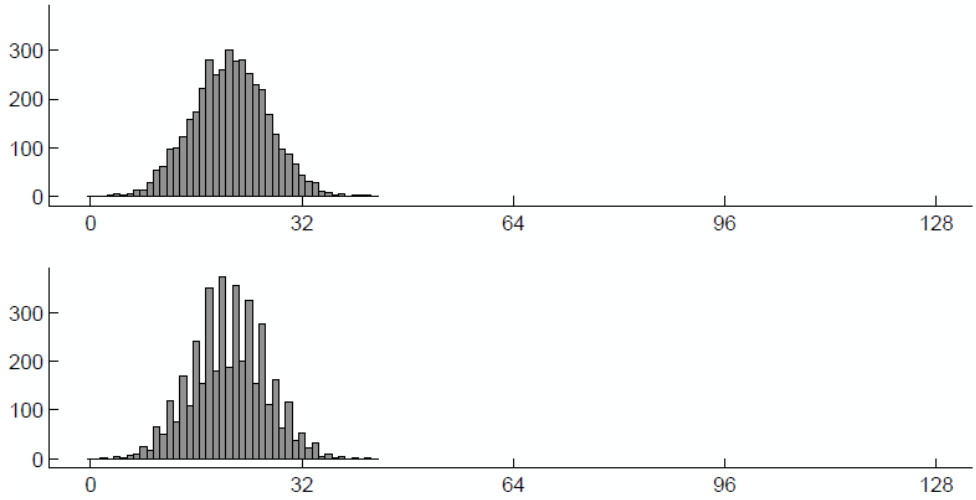
Useimmat tiedostomuotoon perustuvat valokuvien autentikointitekniikat perustuvat JPEG-formaattiin. Tämä on sikäli luonnollista, että JPEG lienee tällä hetkellä ylivoimaisesti suosituin tiedostomuoto valokuvien tallentamiseen. Useimmissa taskukameroissa JPEG on, ellei peräti ainoa mahdollinen, niin ainakin vaihtoehtoinen tallennusmuoto.

JPEG-standardissa määritellään erikseen häviötön ja häviöllinen pakkausmalli. Häviöllisen osan muodostavat *diskreetti kosinimuunnos* (DCT, discrete cosine transformation) ja *kvantisointi*. Diskreetti kosinimuunnos tarkoittaa sitä, että pikseli-informaatio muunnetaan taajuuksiksi, jotka edustavat kuvan erityyppisiä alueita. Tässä yhteydessä kuva jaetaan 8 x 8 pikselin kokoisiin alueisiin, joista muunnos kohdistuu kuhunkin erikseen. Kehikkoa joka jakaa kuvan kyseisiin alueisiin kutsutaan *DCT -kehikoksi*. Muodostuvissa lohkoissa matalat, eli merkitsevimmät taajuudet sijaitsevat vasemmalla ylänurkassa ja korkeat taajuudet alhaalla oikealla. Kvantisointi puolestaan tarkoittaa sitä, että lohkon kertoimet jaetaan ns. *kvantisointimatriisin* luvuilla ja pyöristetään lähimpään kokonaislukuun. JPEG -standardissa on esitetty suositellut matriisit, jotka tarjoavat hyvän tasapainon kuvanlaadun ja pakkauksen välillä. Juuri kvantisoinnin seurauksena osa kuvainformaatiosta menetetään. (Popescu 2005, 70-72)

Popescu ja Farid (2004) ovat havainneet, että kahdesti samaan kuvaan suoritettu kvantisointi aiheuttaa kuvaan eräänlaista jaksoittaista säännönmukaisuutta, joka on havaittavissa kuvan histogrammista. Käytännössä tämä tarkoittaa tasaisin välein esiintyviä huippuja tai notkelmia histogrammin lokeroissa. Vastaavaa ilmiötä ei tavallisesti esiinny vain kertaalleen kvantisoiduissa kuvissa. Tekijöiden ajatus on yksinkertaisesti se, että mikäli kuva havaitaan kahdesti pakatuksi, voi olla syytä epäillä sen aitoutta. Kaksinkertainen pakkaus ei tietenkään yksistään ole erityisen pätevä peruste luokitella kuvaa väärennetyksi. Se voi kuitenkin antaa aihetta tarkemmille tutkimuksille.

Lin ym. (2009) ovat saman ajatuksen pohjalta kehittäneet huomattavasti täsmällisemmän ratkaisun. Lähtökohta on kuitenkin siinä mielessä päinvastainen, että heidän mukaansa ainoastaan kerran kvantisoitujen alueiden voidaan epäillä olevan väärennettyjä, olettaen, että kuvan muut alueet on kvantisoitu kahdesti. On hyvin epätodennäköistä, että muokatuilla alueilla esiintyisi sellaista jaksoittaisuutta, jota kaksoiskvantisointi aiheuttaa. Esimerkiksi pakkaamattomasta kuvasta kopioitu tai ohjelmallisesti generoitu alue ei sisällä lainkaan alkuperäisiä pakkausjälkiä.

Kaksoiskvantisoinnin aiheuttamaa ilmiötä ei myöskään synny, jos muokatun alueen DCT-lohkot eivät ole täysin kohdakkain ensimmäisessä pakkauksessa käytettyjen lohkojen kanssa. Muokatun kuvan histogrammi voidaan ajatella kahden erillisen histogrammin summana. Muokkaamattomat alueet kerryttävät korkeita huippuja, kun taas muokatut alueet jakautuvat satunnaisesti histogrammin lokeroihin. Tällä perusteella voidaan laskea todennäköisyys sille, onko tietty alue muokattu vai ei.




*Piirros 3: Histogrammi normaalijakautuneesta signaalista. Yläkuvassa signaali on kvantisoitu kertaalleen ja alakuvassa kahdesti. (Popescu, 2005, 74)*

Linin ym. (2009) mukaan eräs heidän menetelmänsä merkittäviä vahvuuksia moniin muihin autentikointitekniikoihin verrattuna on laskennallinen nopeus. Yhden kuvan prosessointi kestää vain muutamia sekunteja. Toinen menetelmän vahvuus on se, että se tarkastelu kohdistuu 8 x 8 pikselin kokoisiin alueisiin. Monissa muissa menetelmissä tarkastelu tapahtuu huomattavasti karkeammalla tasolla.

Menetelmän tunnistustarkkuuteen vaikuttaa etenkin JPEG-pakkauksen laatu. Sen vaikutus on kuitenkin erilainen, kuin aiemmin esitellyissä tekniikoissa, sillä tässä tapauksessa tunnistustarkkuus ei heikkene pakkauslaadun heikentyessä. Sen sijaan silloin kun ensimmäisellä ja toisella pakkauskerralla on käytetty täysin samaa laatua, menetelmän tunnistustarkkuus on heikoimmillaan ja putoaa jopa sattuman tasolle. Tästä johtuen menetelmän keskimääräinen tunnistustarkkuus on vain noin 60 prosenttia. Kehittäjien mukaan menetelmän tarkkuutta olisi mahdollista parantaa, mikäli sekä ensimmäisen, että toisen pakkauskerran pakkauslaatu tunnettaisiin. Ongelmaksi muodostuu nimenomaan ensimmäisen pakkauksen taso, sillä sitä ei voida jälkikäteen päätellä suoraan.

Kehittäjien mukaan menetelmä ei toimi siinä tapauksessa, että kuvan DCT -kehikko muuttuu muokkauksen jälkeen. Tämä voi olla seurasta muun muassa koko kuvaan kohdistuvista kierto- tai leikkausoperaatioista, taikka kuvakoon muutoksesta. Yleisesti

ottaen tiedostomuotoon perustuvien tekniikoiden varsin ilmeinen rajoite on niiden soveltuvuus vain yhteen tiedostomuotoon. Linin ym. tekniikka on kuitenkin pienin muutoksin sovellettavissa myös JPEG 2000 -formaattiin.

*Taulukko 1: Yhteenveto passiivisista autentikointitekniikoista*

| Tekniikka | Tunnistetut muokkaustyypit | Vahvuudet | Heikkoudet |
|---|---|---|---|
| Uudelleennäytteistys | Geometriset operaatiot kuten koon muutokset, venytys tai kiertäminen | Muokkauksen tyyppi jossain määrin pääteltävissä | Kaikki uudelleennäytteistyksen asteet eivät ole tunnistettavissa, soveltuu huonosti pakattuihin kuviin |
| Monistetut alueet | Osien kopioiminen kuvan sisällä | Kohtuullisen tarkka myös pakattujen kuvien osalta, muokkaustyyppi tunnetaan | Vain yhden muokkaustyypin tunnistus |
| Interpolaatio | Yleinen | Parhaimmillaan erittäin tarkka | Soveltuu vain kameroille, joissa on värisuodin, ei sovellu JPEG2000-muotoisille kuville |
| PRNU -jälki | Yleinen | Tarkka tunnistus, soveltuu myös kameran yksilöimiseen | Vaatii useita muokkaamattomia kuvia, hitaus, tummat monimutkaiset alueet ongelmallisia |
| Kohinan tason muutokset | Yleinen | Soveltuu kohtuullisen hyvin pakattujen kuvien käsittelyyn, jos käytetään suurta lohkokokoa | Kohinan taso voi vaihdella aidoissakin kuvissa |
| JPEG kaksoiskvantisaatio | Yleinen | Nopeus, alueellinen tarkkuus, soveltuu hyvin pakattujen kuvien käyttöön | Rajoittuu JPEG-formaattiin, ei toimi jos molemmilla pakkauskerroilla käytetty samaa laatua |

# 4 YHTEENVETO

Vesileimatekniikoiden avulla digitaalisten valokuvien aitous voidaan varmistaa suhteellisen luotettavasti ja tarkasti. Parhaimmat toteutukset kykenevät paikallistamaan muokatut alueet jopa pikselin tarkkuudella. Vesileimatekniikat kykenevät myös tunnistamaan periaatteessa minkä tyyppisiä muokkauksia hyvänsä. Useat toteutukset kykenevät myös erottamaan kuvan merkitystä muuttavat varsinaiset muokkaustoimet ja merkityksen säilyttävät muutokset, kuten häviöllisen pakkauksen. Vesileimatekniikoiden käyttöalaa rajoittaa merkittävästi se, että ne eivät sovellu muiden, kuin vesileimalla varustettujen kuvien autentikointiin. Kuitenkin suurin osa tällä hetkellä olemassa olevista ja kaiken aikaa syntyvistä kuvista on vesileimattomia. Lisäksi vesileiman lisäämisen tulisi tapahtua jo kuvan ottamisen hetkellä, jotta valokuvan aitous voitaisiin varmasti taata. Ainakin toistaiseksi tämä on varsin harvinainen mahdollisuus markkinoilla olevissa kameroissa. Luultavasti täyttä varmuutta ei ole myöskään siitä, ovatko vesileimat riittävän turvallisia, jotta niitä ei olisi missään olosuhteissa mahdollista väärentää tai muuttaa luvattomasti.

Optimaalisissa olosuhteissa valokuvien autentikointi onnistuu myös erilaisilla passiivisilla tekniikoilla varsin luotettavasti ja tarkasti. Passiivisten tekniikoiden käyttöön liittyy kuitenkin hyvin paljon epävarmuustekijöitä, sillä olosuhteet eivät aina ole lähimainkaan optimaalisia. Esimerkiksi kuvien häviöllinen pakkaaminen, etenkin runsaasti käytettynä, on useimpien passiivisten tekniikoiden kannalta enemmän tai vähemmän ongelmallista. Lisäksi jotkin kamerat kohdistavat kuviin melko voimakasta jälkiprosessointia, kuten kohinan poistoa. Lienee jossain määrin epäselvää, miten tällaiset toimenpiteet vaikuttavat passiivisten autentikointitekniikoiden toimintaan. Osa tekniikoista on myös käyttöalaltaan melko rajallisia. Kulloisetkin olosuhteet määrittelevät varmasti eri tekniikoiden käyttöarvon. Esimerkiksi osa tekniikoista soveltuu toisia paremmin pakattujen kuvien käsittelyyn. Joissain tilanteissa autentikointimenetelmän nopeudella voi olla kriittinen merkitys, vaikkapa käsiteltävien kuvien suuren lukumäärän vuoksi. Toisessa tilanteessa puolestaan tunnistuksen tarkkuus saattaa olla oleellista. Kuten Chen ym. (2007) toteavat, väärennettyjen valokuvien tunnistaminen on monimutkainen ongelma, johon ei ole yhtä universaalia ratkaisua. Luotettavan päätöksen tekemiseen tarvitaan erilaisia menetelmiä. Tässä tutkielmassa käsitellyt tekniikat on valittu osin sillä perusteella, että ne voisivat mahdollisimman hyvin kompensoida toistensa puutteita.

Parhaiten passiiviset tekniikat toimivat silloin, kun väärentäjä ei niitä tunne, eikä näin ollen osaa peittää jälkiään oikealla tavalla. Toisaalta yrittäessään peittää ihmissilmällä havaittavia muokkausjälkiä, tietämätön väärentäjä saattaa jättää sellaisia jälkiä, jotka on helppo havaita edellä esiteltyjen tekniikoiden avulla. Useimpia menetelmiä varten lienee kuitenkin mahdollista kehittää vastatekniikoita, ja joitakin menetelmiä vastaan sellaisia on jo kehitetty. Tosin ainakin tällä hetkellä vastatekniikoiden käyttö vaatii siinä määrin ammattitaitoa, etteivät ne ole välttämättä amatööriväärentäjien käytettävissä. Myös vastatekniikoita ajatellen useiden eri autentikointimenetelmien käyttö voi olla hyödyllistä. Gloe ym. (2007) toteavatkin, että on huomattavasti hankalampaa kehittää yleisluontoinen vastahyökkäys kaikkia mahdollisia autentikointitekniikoita vastaan, kuin yhtä nimenomaista tekniikkaa vastaan.

Sen paremmin aktiivisten kuin passiivistenkaan autentikointitekniikoiden avulla ei näyttäisi olevan kovinkaan helppoa päätellä kuvaan kohdistetun muokkauksen tyyppiä. Päätelmiä voidaan toki tehdä ainakin sellaisten menetelmien avulla, jotka rajoittuvat tiettyjen muokkaustyyppien tunnistamiseen. Sen sijaan tekniikat, jotka pyrkivät tunnistamaan kaikki mahdolliset kuvaan kohdistuvat muutokset, eivät yleensä kykene erottelemaan minkä tyyppisestä muutoksesta on kyse. Poikkeuksen muodostavat sellaiset vesileimatekniikat, jotka karkealla tasolla kykenevät palauttamaan muokatun alueen alkuperäisen sisällön. Passiivisten tekniikoiden osalta on myös osin epäselvää, mitkä muokkaustyypit ovat ylipäätänsä mahdollista havaita.

Digitaalisten valokuvien aitouden arviointi on toistaiseksi varsin tuore tutkimusala. Se on kuitenkin herättänyt viime aikoina paljon kiinnostusta, joten entistä parempia menetelmiä syntyy varmasti jatkossakin. Toisaalta olemassa olevia menetelmiä vastaan kehitetään luultavasti entistä tehokkaampia vastatoimia. Farid (2009) toteaakin, että valokuvien autentikointitekniikat voivat tehdä uskottavien väärennösten luomisen entistä vaikeammaksi ja hitaammaksi, mutta tuskin koskaan täysin mahdottomaksi.

Mahdollisia jatkotutkimuksen aiheita voisivat olla ainakin uusien autentikointimenetelmien kehittäminen, olemassa olevien menetelmien parantaminen tai eri menetelmien yhdistäminen. Tämän tutkielman tuloksia olisi varsin luonnollista käyttää pohjana etenkin viimeksi mainittuun tarkoitukseen. Eri menetelmien käytännön toteutus laboratorio-olosuhteissa voisi tarjota konkreettista näyttöä muun muassa siitä, kuinka tarkkoihin tuloksiin menetelmiä yhdistämällä päästään. Tällöin tarvittaisiin mahdollisimman kattava testiaineisto, joka sisältäisi laadultaan ja sisällöltään vaihtelevia, eri tavoin muokattuja kuvia. Kokeellinen tutkimus saattaisi myös paljastaa erilaisia tekijöitä, joilla voi olla vaikutusta menetelmien yhteiskäyttöön. Esimerkiksi yhteenlaskettu laskenta-aika saattaisi olla jossain tilanteessa käytännön näkökulmasta epärealistinen. Toisaalta olisi varmasti hyödyllistä miettiä sitä, voidaanko laskennallisia kustannuksia tällöin pienentää, esimerkiksi etsimällä eri algoritmeista vaiheita, joita voitaisiin mahdollisesti yhdistää. Laskennallisia kustannuksia voitaisiin mahdollisesti pienentää myös arvioimalla sitä, mitä tekniikoita tietyssä tilanteessa kannattaa ylipäätänsä käyttää, ja missä järjestyksessä. Esimerkiksi jos kuva havaitaan voimakkaasti pakatuksi, ei joidenkin autentikointimenetelmien käyttö luultavasti hyödytä mitään. Hyödyllistä voisi olla myös tutkia sitä, missä määrin erilaiset muokkaustyypit ovat tunnistettavissa.

# LÄHDELUETTELO

Chen, M., Fridrich, J., Lukáš, J., & Goljan, M. 2007. Imaging sensor noise as digital X-ray for revealing forgeries. Teoksessa Information hiding, Lecture Notes in Computer Science. Saint Malo, France: Springer Berlin / Heidelberg, 342-358.

Chun-Shien Lu & Liao, H. -. M. 2001. Multipurpose watermarking for image authentication and protection. Image Processing, IEEE Transactions on 10(10), 1579-1592.

Cook, D. W., & Rajan, P. K. 2006. TIAMAT: A new fragile watermarking technique for image authentication. Teoksessa System Theory, 2006. SSST '06. Proceeding of the Thirty-Eighth Southeastern Symposium on, March 5-7. Cookeville, TN, USA: 221-225.

Cox, I. J., & Miller, M. L. 1997. Review of watermarking and the importance of perceptual modeling. Teoksessa Bernice E. Rogowitz & Thrasyvoulos N. Pappas (toim.) Human Vision and Electronic Imaging II, Perception and Watermarking, February 10-13. San Jose, CA, USA: SPIE, 92-99.

Farid, H. 2009. Image forgery detection. Signal Processing Magazine, IEEE 26(2), 16-25.

Fridrich, J. 1998. Image watermarking for tamper detection. Teoksessa Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, Chicago, IL, USA, October 4-7. Chicago, IL, USA: IEEE, 404-408.

Fridrich, J. 2002. Security of fragile authentication watermarks with localization. Teoksessa Edward J. Delp III & Ping W. Wong (toim.) Proc. SPIE Photonic West, Electronic Imaging 2002, Security and Watermarking of Multimedia Contents, April 29, 2002. San Jose, California: SPIE, 691-700.

Fridrich, J. 2009. Digital image forensics. Signal Processing Magazine, IEEE 26(2), 26-37.

Gallagher, A. C., & Tsuhan Chen. 2008. Image authentication by detecting traces of demosaicing. Teoksessa Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on, June 23-28. Los Alamitos, CA, USA: IEEE Computer Society, 1-8.

Gloe, T., Kirchner, M., Winkler, A., & Böhme, R. 2007. Can we trust digital image forensics? Teoksessa Rainer Lienhart, Anand R. Prasad, Alan Hanjalic, Sunghyun Choi, Brian P. Bailey & Nicu Sebe (toim.) MULTIMEDIA '07: Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany, September 24-29. New York, NY, USA: ACM Press, 78-86.

Hameed, K., Mumtaz, A., & Gilani, S. A. M. 2006. Digital image watermarking in the wavelet transform domain. Teoksessa Proceedings of World Academy of Science, Engineering and Technology, Helsinki, Finland, January, 2006. WASET, 13(2006)

Hartung, F. & Kutter, M. 1999. Multimedia watermarking techniques. Proceedings of the IEEE 87(7), 1079-1107.

Holliman, M. & Memon, N. 2000. Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes. Image Processing, IEEE Transactions on 9(3), 432-441.

Kundur, D. & Hatzinakos, D. 1999. Digital watermarking for telltale tamper proofing and authentication. Proceedings of the IEEE 87(7), 1167-1180.

Kutter, M., & Petitcolas, F. A. P. 1999. A fair benchmark for image watermarking systems. Teoksessa Ping Wah Wong & Edward J. Delp (toim.) Electronic Imaging '99, Security and Watermarking of Multimedia Contents, January 25-27. San Jose, California, USA: I.S.&T. and S.P.I.E., 226-239.

Li, C. T. 2004. Digital fragile watermarking scheme for authentication of JPEG images. Vision, Image and Signal Processing, IEE Proceedings - 151(6), 460-466.

Lin, E. T., Podilchuk, C. I., & Delp III, E. J. 2000. Detection of image alterations using semifragile watermarks. Teoksessa Ping W. Wong & Edward J. Delp III (toim.) Security and Watermarking of Multimedia Contents II, January 24. San Jose, CA, USA: SPIE, 152-163.

Lin, Z., He, J., Tang, X. & Tang, C. 2009. Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis. Pattern Recognition 42(11), 2492-2501.

Mahdian, B. & Saic, S. 2009. Using noise inconsistencies for blind image forensics. Image and Vision Computing 27(10), 1497-1503.

Memon, N., & Fridrich, J. 2000. Further attacks on the yeung-mintzer fragile watermark. Teoksessa Ping W. Wong & Edward J. Delp III (toim.) Proceedings of IS&T/SPIE's 12th Annual Symposium, Electronic Imaging 2000: Security and Watermarking of Multimedia Content II, January 24. San Jose, CA, USA: SPIE, 428-437.

Miller, M. L., Cox, I. J., Linnartz, J. M. G., & Kalker, T. 1999. A review of of watermarking principles and practices. Teoksessa K. K. Parhi & T. Nishitani (toim.) Digital signal processing for multimedia systems. New York, NY, USA: IEEE, 461-485.

Mohanty, S. P. 1999. Digital Watermarking : A Tutorial Review. Dept of Computer Science and Engineering, Unversity of South Florida.

Podilchuk, C. I. & Delp, E. J. 2001. Digital watermarking: Algorithms and applications. Signal Processing Magazine, IEEE 18(4), 33-46.

Podilchuk, C. I. & Wenjun Zeng. 1998. Image-adaptive watermarking using visual models. Selected Areas in Communications, IEEE Journal on 16(4), 525-539.

Popescu, A. C. 2005. Statistical Tools for Digital Image Forensics. Department of Computer Science, Darthmouth College, Ph.D. Dissertation.

Popescu, A. C., & Farid, H. 2004. Statistical tools for digital forensics. Teoksessa 6th International Workshop on Information Hiding, Toronto, Canada, May. Toronto, Canada: Springer, 128-147.

Popescu, A. C. & Farid, H. 2005. Exposing digital forgeries in color filter array interpolated images. Signal Processing, IEEE Transactions on 53(10), 3948-3959.

Potdar, V. M., Han, S., & Chang, E. 2005. A survey of digital image watermarking techniques. Teoksessa Industrial Informatics, 2005. INDIN '05. 2005 3rd IEEE International Conference on, August 10-12. 709-716.

Rey, C. & Dugelay, J. 2002. A survey of watermarking algorithms for image authentication. EURASIP Journal on Applied Signal Processing 2002(6), 613-621.

Servetto, S. D., Podilchuk, C. I., & Ramchandran, K. 1998. Capacity issues in digital image watermarking. Teoksessa Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, Chicago, IL, USA, October 4-7. Los Alamitos, CA, USA: IEEE Computer Society, 445-449.

Skodras, A., Christopoulos, C. & Ebrahimi, T. 2001. The JPEG 2000 still image compression standard. Signal Processing Magazine, IEEE 18(5), 36-58.

Wolfgang, R. B., Podilchuk, C. I. & Delp, E. J. 1999. Perceptual watermarks for digital images and video. Proceedings of the IEEE 87(7), 1108-1126.

Wong, P. W. 1998. A watermark for image integrity and ownership verificatio. Teoksessa PICS 1998: IS&T's 1998 Image Processing, Image Quality, Image Capture, Systems Conference, May 17-20. Portland, Oregon, USA: IS&T - The Society for Imaging Science and Technology, 374-379.

Yong-Zhong He, & Zhen Han. 2008. A fragile watermarking scheme with pixel-wise alteration localisation. Teoksessa Signal Processing, 2008. ICSP 2008. 9th International Conference on, Beijing, China, October 26-29. 2201-2204.

Yunfei Jiang, & Ping Guo. 2007. Comparative studies of feature extraction methods with application to face recognition. Teoksessa Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on, Montreal, Quebec, Canada, October 7-10. 3627-3632.

Zhang, Z., Ren, Y., Ping, X.-J., He, Z.-Y. & Zhang, S.-Z. 2008. A survey on passive-blind image forgery by doctor method detection.