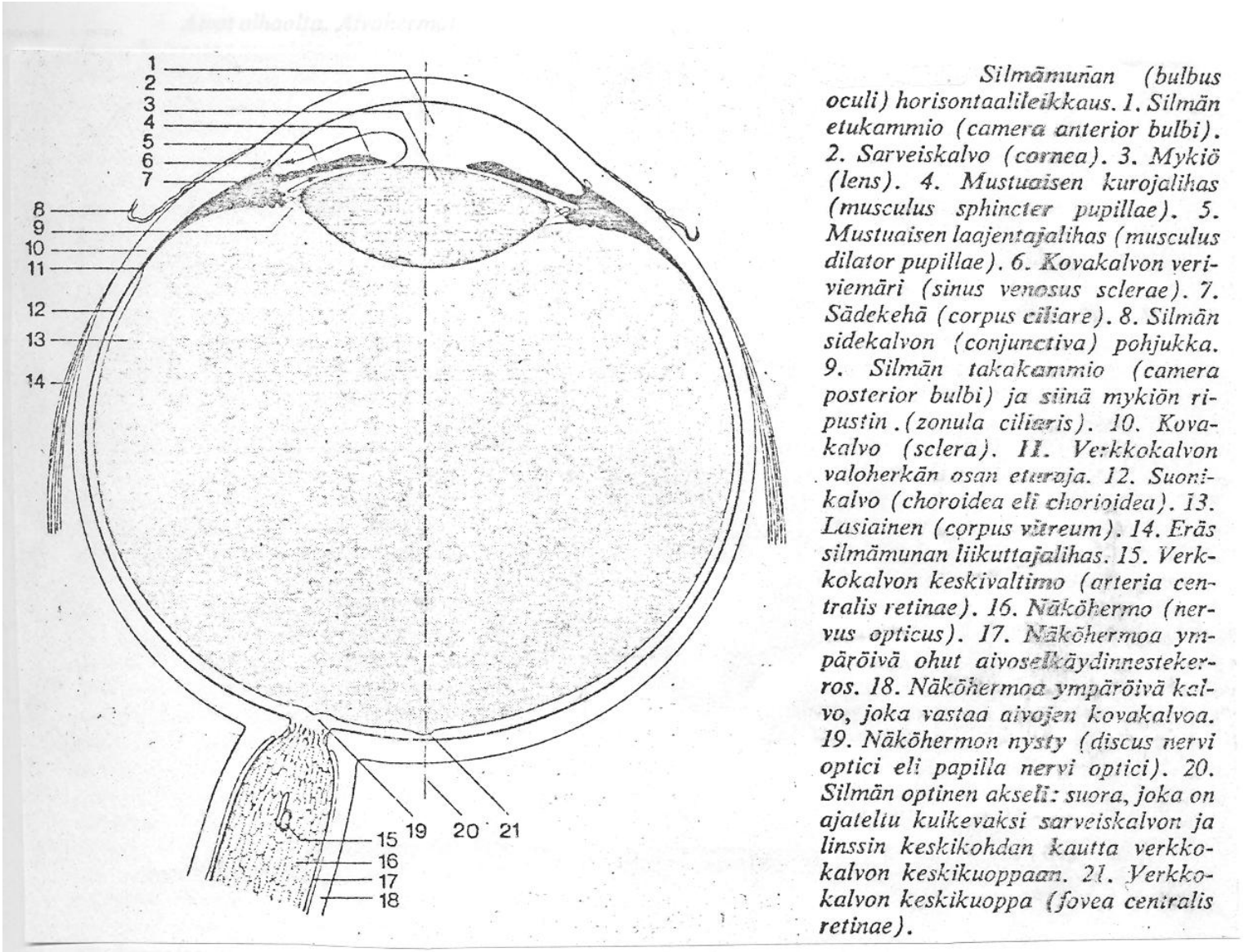
2. Digitaalisten kuvien peruskäsitteet

2.1. Visuaalinen havaitseminen

Tässä luvussa käsitellään digitaalisten kuvien perusasioista, aina kuvien näkemisestä pikseleihin ja tarvittaviin laskentamenetelmiin asti.

Vaikka kuvanprosessointi perustuu matemaattisten kaavojen käyttöön, inhimillisellä intuitiolla ja analysointikyvyllä on keskeinen rooli eri menetelmien valinnassa. Monesti on kiinnostava lopputuote, miltä kuva ”näyttää”, kun taas toisinaan on olennaista tunnistaa kuvasta joitakin kiinnostavia kohteita. Silmäillään seuraavassa suppeasti ihmissilmän rakennetta ja myös näkökyvyn rajoituksia, jotka on otettava huomioon digitaalisten kuvien käsittelyssä.

Kuva 2.1. sisältää silmän poikkileikkauksen. Silmien väli on aikuisella noin 6 cm ja silmän keskimääräinen halkaisija 20 mm. Silmä on geometrisesti lähes pallo.



Silmämunan (bulbus oculi) horisontaalileikkaus. 1. Silmän etukammio (camera anterior bulbi). 2. Sarveiskalvo (cornea). 3. Mykiö (lens). 4. Mustuaisen kurojalihas (musculus sphincter pupillae). 5. Mustuaisen laajentajalihas (musculus dilator pupillae). 6. Kovakalvon veriviemäri (sinus venosus sclerae). 7. Sädekehä (corpus ciliare). 8. Silmän sidekalvon (conjunctiva) pohjukka. 9. Silmän takakammio (camera posterior bulbi) ja siinä mykiön ripustin (zonula ciliaris). 10. Kovakalvo (sclera). 11. Verkkokalvon valoherkän osan eturaja. 12. Suonikalvo (choroidea eli chorioidea). 13. Lasiainen (corpus vitreum). 14. Eräs silmämunan liikuttajalihas. 15. Verkkokalvon keskivaltimo (arteria centralis retinae). 16. Näköhermo (nervus opticus). 17. Näköhermoa ympäröivä ohut aivoselkäydinnestekerros. 18. Näköhermoa ympäröivä kalvo, joka vastaa aivojen kovakalvoa. 19. Näköhermon nysty (discus nervi optici eli papilla nervi optici). 20. Silmän optinen akseli: suora, joka on ajateltu kulkevaksi sarveiskalvon ja linssin keskikohdan kautta verkkokalvon keskikuoppaan. 21. Verkkokalvon keskikuoppa (fovea centralis retinae).

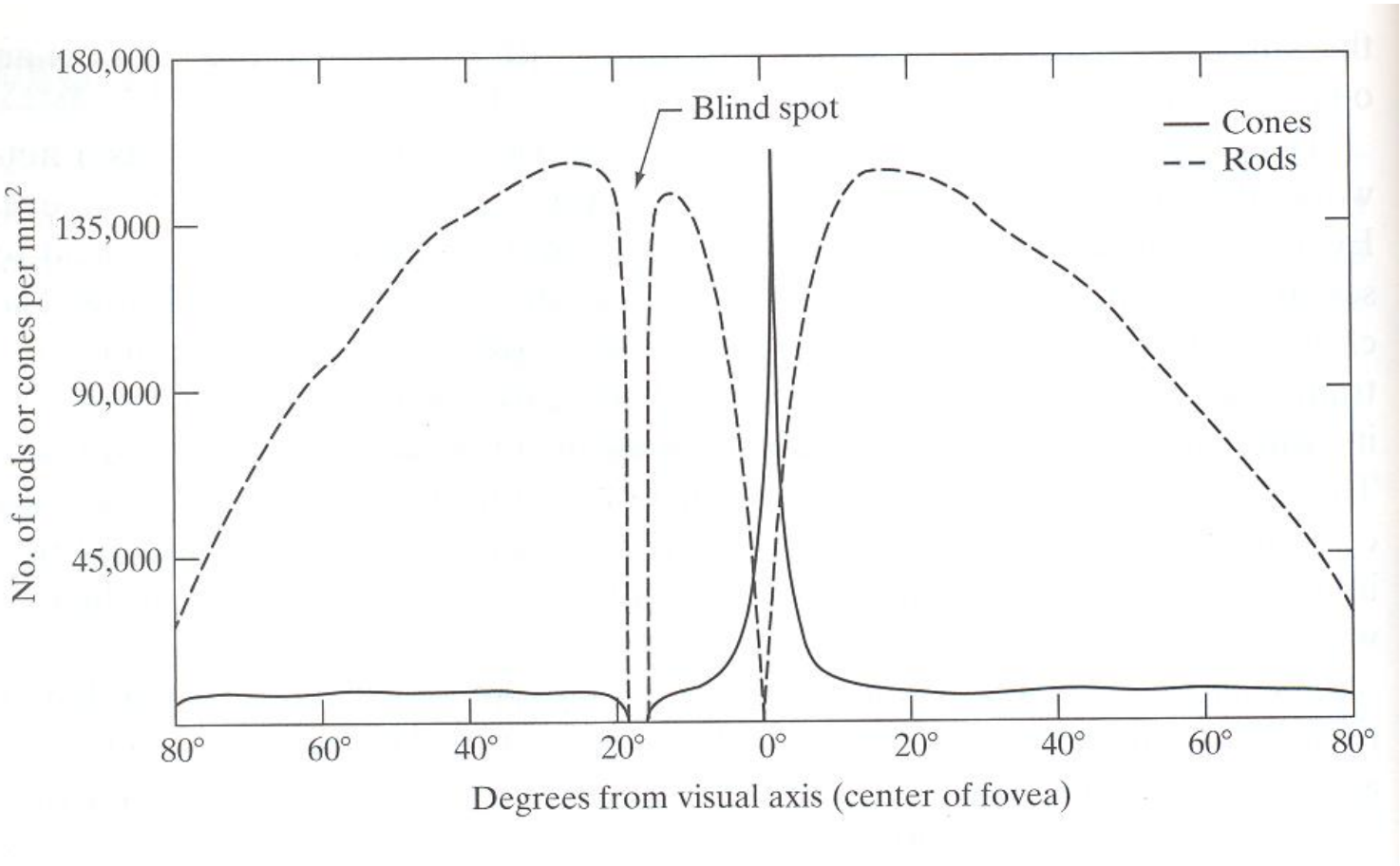
Kuva 2.1. Silmän osat.

Cornea eli sarveiskalvo on ulommainen kalvo silmän päällä. Iiris eli mustuainen ympäröi linssiä eli mykiötä, jonka läpi valo kulkee nesteen läpi retinalle eli verkkokalvolle. Iiris laajenee päästääkseen hämärässä lisää valoa silmään ja suppenee kirkkaassa valossa. Tarkan näön alue on pinta-alaltaan pieni fovea eli verkkokalvon keskikuoppa. Katsetta siirrettäessä kohteeseen kuva muodostuu fovealle.

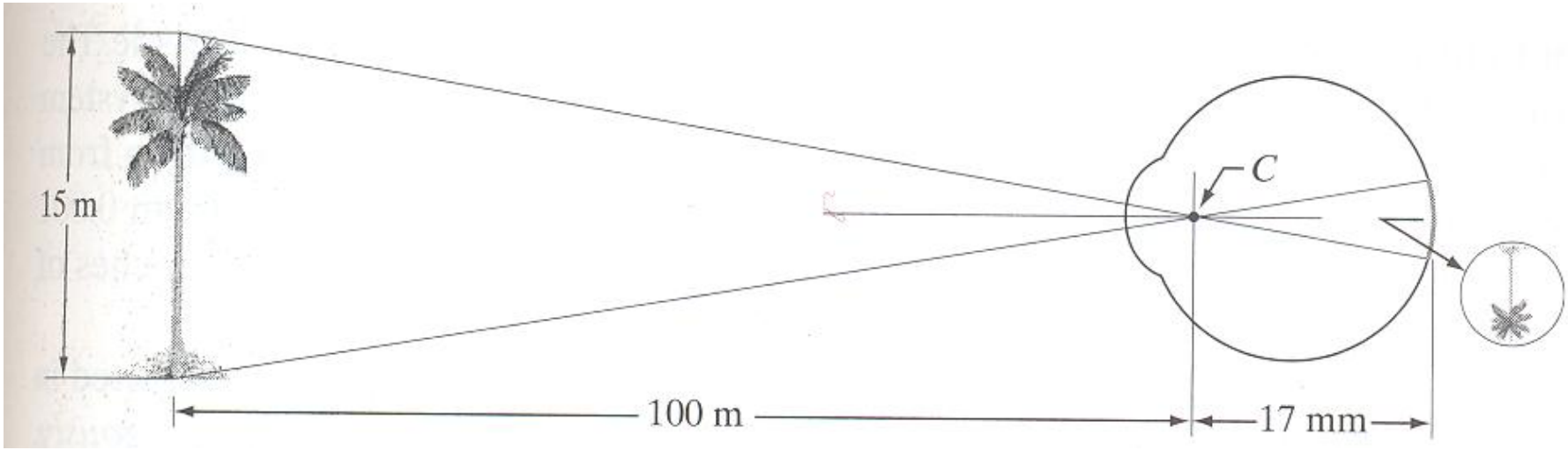
Näkökentässä olevasta kohteesta heijastuu valoa retinalle, kun silmä on sopivasti fokusoitunut, ts. mykiö on säädetty sopivasti kohdentumaan katsottavaan kohteeseen. Mykiö absorboi noin 8 % näkyvästä valosta ja erityisesti näkyvän valon ”reunoilta” infrapuna- ja ultraviolettisäteitä, jotka suuressa määrin absorboituessa voivat tämän vuoksi jopa vahingoittaa silmää

Verkkokalvolla näköaistimuksen ”anturina” ovat 6-7 miljoonaa tappisolua ja 75-150 miljoonaa sauvasolua. Edelliset ovat keskittyneet foveaan ja aistivat värejä. Jälkimmäiset ovat herkkiä valolle. Silmän kuusi liikuttajalihasta pyörittävät silmämunaa siirrettäessä katsetta. Silmämuna liikkuu ikään kuin kolmen akselin, (kuviteltujen) vaaka-, pysty- ja kiertoakselinsa suunnassa. Kiertoakselin voidaan ajatella yhtyvän kuvaan 2.1. piirrettyyn optiseen näköakseliin.

Päätä liikuttamatta ihmisen katselukulma on runsas ±80° (kuva 2.2.) Silmät ovat kuitenkin vain aistinelin, ja varsinainen informaation prosessointi tapahtuu aivoissa, jonne näköhermot välittävät aistitun informaation ja jossa kirjaimellisesti tehdään näkemiseen liittyvä ”aivotyö”. Kuva 2.3. esittää, miten kuva muodostuu verkkokalvolle.



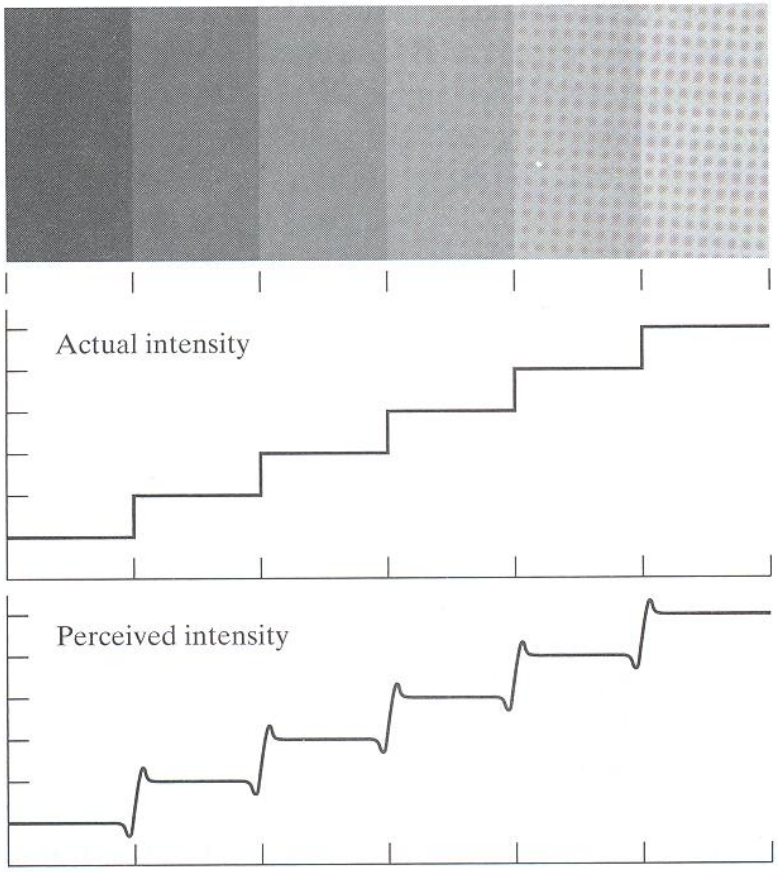
Kuva 2.2. Sauva- ja tappisolujen jakauma verkkokalvolla.



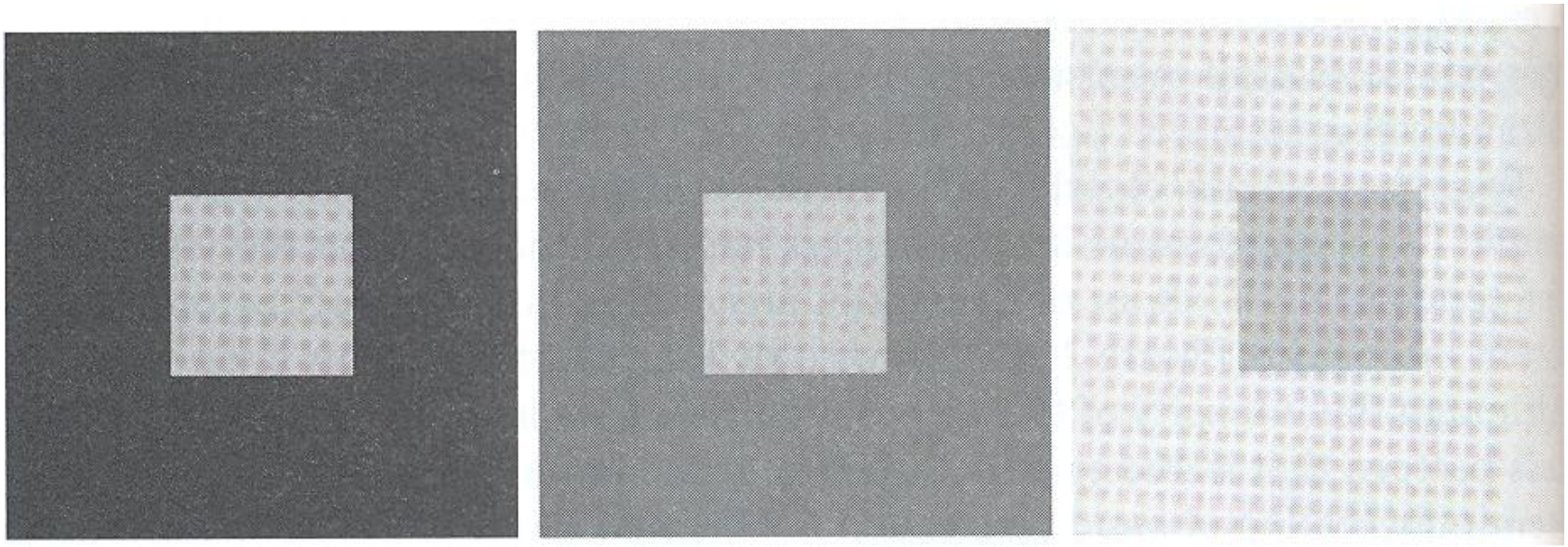
Kuva 2.3. Kaavakuva kuvan muodostumisesta verkkokalvolle linssin optisen keskipisteen *C* kautta. Kuvan korkeus on suhteessa 15/100=*h*/17, josta saadaan *h*=2.55 mm. Kun kuva on näin piskuinen, on selvää, että aistinsoluja tarvitaan suurella tiheydellä verkkokalvolla.

Kun digitaaliset kuvat esitetään intensiteettitasojen diskreetteinä joukkoina, silmän kyky erottaa eri intensiteettien välillä on tärkeä ominaisuus esitettäessä kuvanprosessoinnin tuloksia. Silmä pystyy adaptoitumaan peräti luokkaa 10^{10} olevassa suhteessa. Kokeelliset tutkimukset ovat osoittaneet *subjektiivisen* kirkkauden (kuten tämä on havaittu) olevan intensiteetin logaritminen funktio.

Kirkkauden havaitseminen ei kuitenkaan ole intensiteetin yksinkertainen funktio. Visuaalinen järjestelmä on taipuvainen alittamaan tai ylittämään eri intensiteettirajat. Kuva 2.4. demonstroi tätä. Nämä *Machin kaistat* havaitsi ensimmäisen Ernst Mach 1865. Toinen ilmiö on *samanaikaiset kontrastit*, kun alueen havaittu kirkkaus ei riipu suoraviivaisesti intensiteetistään, kuten kuva 2.5. osoittaa.

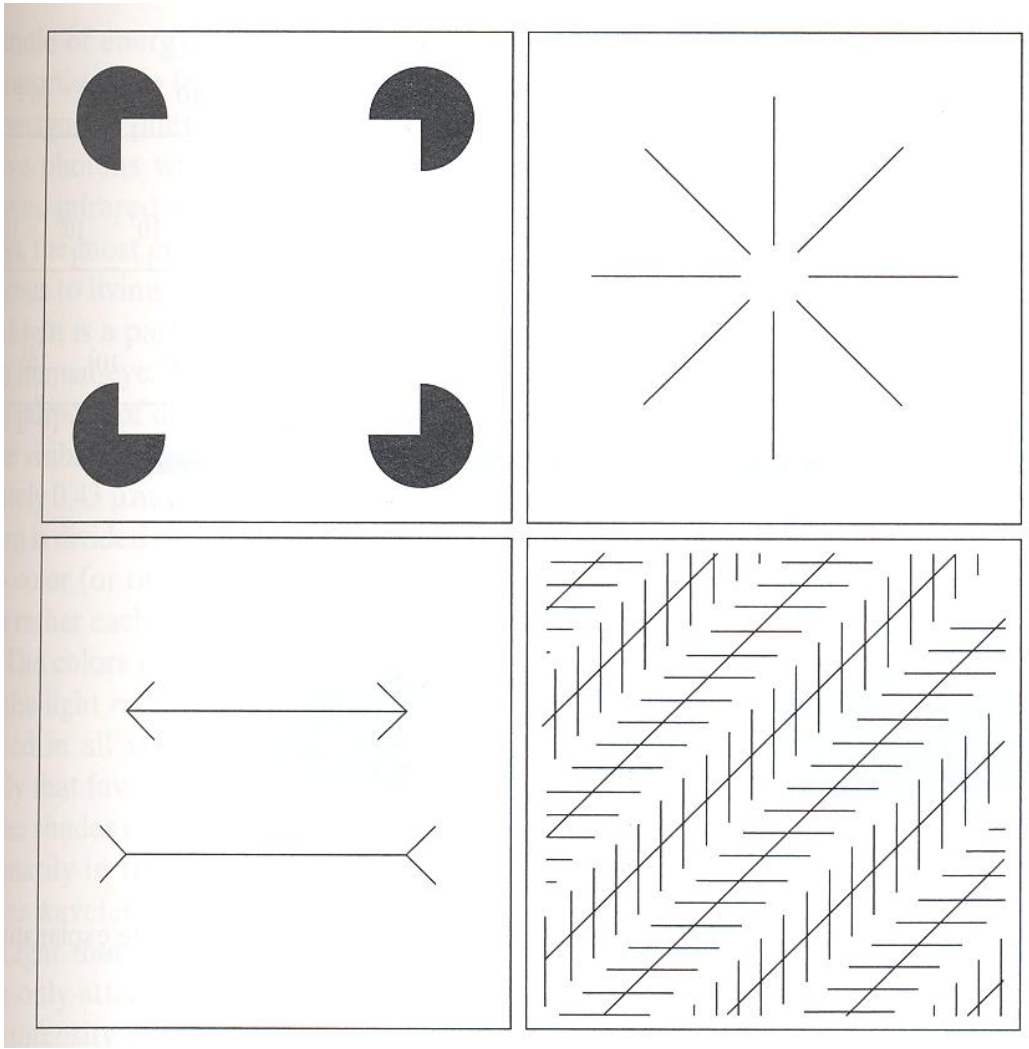


Kuva 2.4. Vaikka kaistojen intensiteetti on kullakin vakio, havaitaan kirkkaus ikään kuin pykäläisenä rajan kohdalla.



Kuva 2.5. Samanaikainen kontrasti: havaittu kirkkaus ei riipu vain intensiteetistä. Kaikilla keskineliöillä on täsmälleen sama intensiteetti, mutta ne näyttävät tummenevan taustan vaaletessa. Vastaava ilmiö nähdään katsottaessa ensin valkoista paperia pöydällä ja toiseksi nostettaessa se ylös kirkasta taivasta vasten.

Muita näköhavaintojen ilmiöitä ovat optiset illuusiot, joissa nähdään olematonta informaatiota tai geometrisesti virheellisesti kohteiden ominaisuuksia. Kuva 2.6. esittää esimerkkejä. Sen vasemmassa yläosassa voidaan nähdä neliö, vaikka siinä ei ole mitään neliön ääriviivoja. Vastaavasti voidaan nähdä ympyrä kuva oikeassa yläosassa. Huomaa, miten vähäinen määrä suoria riittää antamaan illuusion niiden keskellä olevasta ympyrästä. Vasemman alaosan vaakasuorat janat ovat tarkasti yhtä pitkät, vaikka erilaiset kärkikuviot harhauttavat näyttämään ne eripituisiksi. Oikean alaosan suorat näyttävät olevan erisuuntaisia, mutta eivät todellisuudessa ole, vaan monet lyhyet samansuuntaiset poikkiviivat aiheuttavat illuusion.



Kuva 2.6. Optisia illuusioita.

2.2. Valo ja sähkömagneettinen spektri

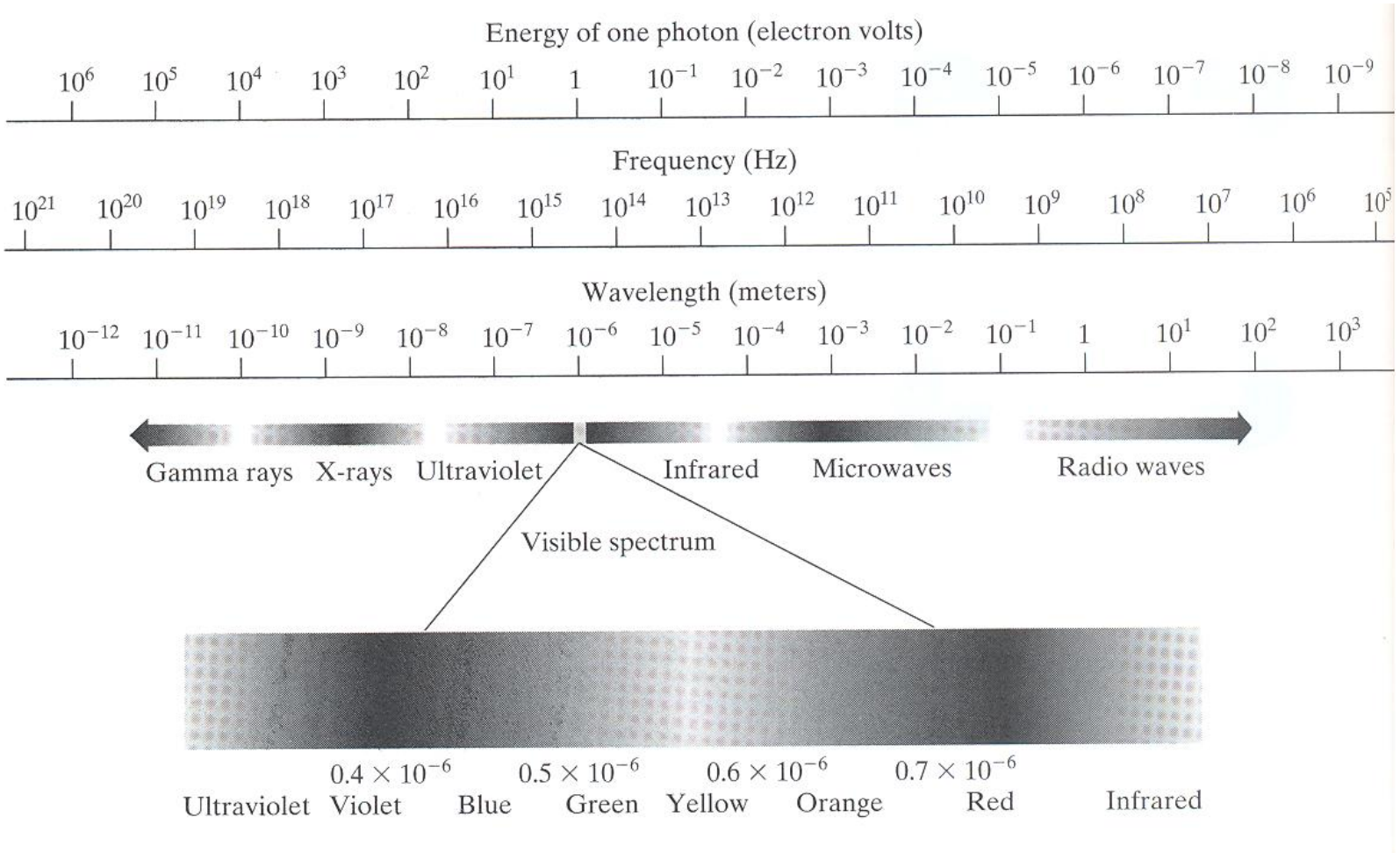
Isaac Newton havaitsi 1666 valkoisen valon olevan todellisuudessa yhdistelmä kaikkia mahdollisia (silmin nähtäviä) värejä kokeillessaan valon hajaantumista lasiprisman läpi. Valo muodostuu värien jatkuvasta spektristä ulottuen violetista punaiseen. Näkyvä valo edustaa ainoastaan sähkömagneettisen spektrin kapeaa kaistaa (kuva 2.7.). Kaikkiaan spektri ulottuu gammasäteistä radioaaltoihin. Nämä ilmaistaan aallonpituuden λ, taajuuden f ja valon nopeuden c=2.988·10^8 m/s perusteella

$$\lambda = \frac{c}{f}. \quad (1)$$

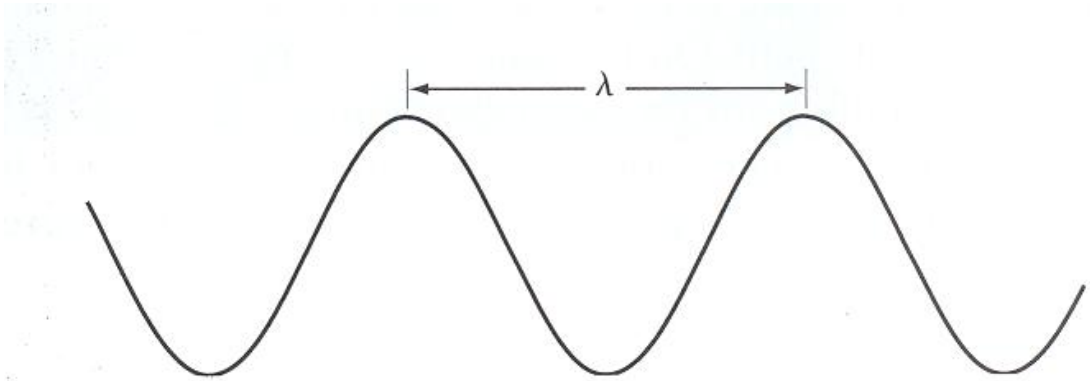
Energia (fotoni) tässä yhteydessä on

$$E = hf, \quad (2)$$

jossa h=6.62·10^{-34} Js on Planckin vakio. Aaltoliikettä kuvataan kuvan 2.8. tapaan. Ihmisen silmä havaitsee herkimmin oranssia (λ=5550 Å, 1 Å=10^{-10} m).



Kuva 2.7. Sähkömagneettinen spektri, josta näkyvän valon kaista on varsin kapea osa.



Kuva 2.8. Aallonpituus graafisena esityksenä.

Sähkömagneettisilla aalloilla on kaksinainen luonne: kuvan 2.8. kaltainen siniaaltoliike tai se voidaan ymmärtää massattomien hiukkasten aaltoliikkeenä, joka etenee valon nopeudella. Jokaisen hiukkasen energia on fotonin suuruinen. Kaavan (2) mukaan energia on suhteessa taajuuteen. Mikroaalloilla on enemmän energiaa kuin radioaalloilla ja kaikkein eniten on gammasäteilyllä, joka on täten vaarallista eläville organismeille.

Ihmisen silmä kykenee aistimaan määrättyä sähkömagneettista säteilyä, valoa, jonka aallonpituus on violetin 0.43 µm:stä punaisen 0.79 µm:iin. Käytännöllisyyden vuoksi nämä aallonpituudet jaetaan kuuteen väriin: violetti, sininen, vihreä, keltainen, oranssi ja punainen. Niillä ei ole mitään tarkkaa rajaa, vaan pikemmin muuttuvat tasaisesti väristä toiseen.

Ihmisen kohteesta näkemät värit riippuvat siitä heijastuneesta valosta. Kappale, joka heijastaa melko tasaisesti kaikilla valon aallonpituuksilla, nähdään valkoisena. Sen sijaan kappale, joka heijastaa vain joitakin värispektrin osia, nähdään vastaavissa väreissä. Esim. vihreä näkyy aallonpituuksina 500-570 nm.

Monokromaattisen eli yksivärisen valon ainoa ominaisuus on sen intensiteetti eli valovoima. Harmaasävy vastaa termeissämme tätä, kun ei ole värikuvista kyse.

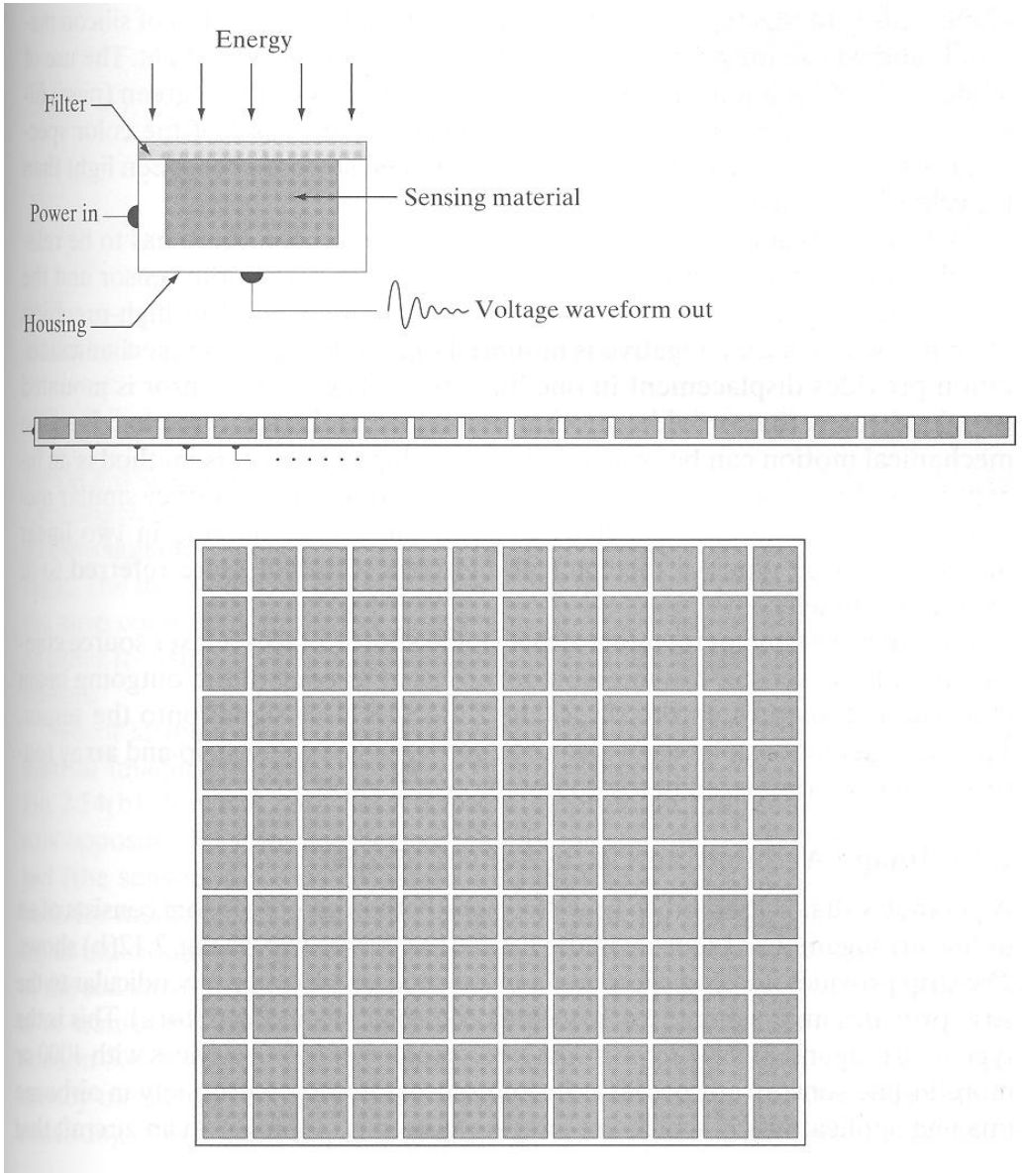
Kromaattinen eli värillinen valo käsittää joitakin osia valon spektristä. Taajuuden eli frekvenssin lisäksi sitä määritetään kolmella suureella: radianssi, luminanssi eli valovirta ja kirkkaus. Radianssi [watti, W] tarkoittaa energian määrää, joka tulee valolähteestä. Luminanssi [lumen, lm] mittaa havaitsijan havaitsemaa valolähteen energiaa. Kirkkaus on subjektiivinen suure.

2.3. Kuvan hankinta

Valaistus ja näkymä kolmiulotteisessa arkimaailmassamme ovat luonnollisesti ensimmäinen mieleen tuleva kuvauskohde. Digitaalinen kuvaaminen käsittää kuitenkin monia teknisiä sovellusaloja ja lääketieteellisen kuvantamisen, joista esimerkkinä on röntgenkuvat.

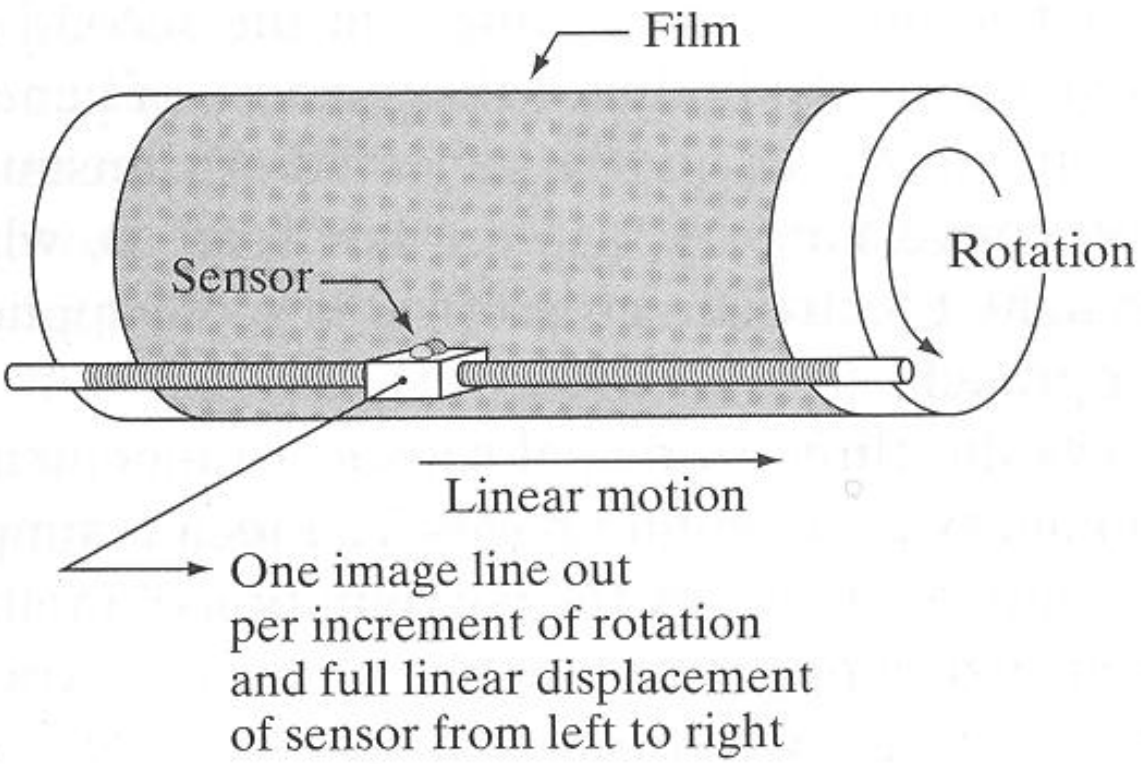
Kuva 2.9. esittää kolme pääasiallista tapaa järjestää kuvausanturit tai -sensorit. Niihin tuleva energia muutetaan jännitteeksi sähkölähteen ja anturimateriaalin avulla. Jälkimmäinen on herkkä määrätynlaiselle energialle, jota halutaan havaita. Anturien vaste on ulostulojännitteen aaltosignaali. Tämä digitoimalla eli kvantisoimalla kustakin anturista saadaan digitaalinen kuva. Yksittäisiä antureja tarvitaan luonnollisesti suuri määrä kuvaa varten (kuva 2.9.).

Kuva 2.9. Ylinnä yksittäinen anturi, keskellä anturiliuska ja alhaalla anturimatriisi.

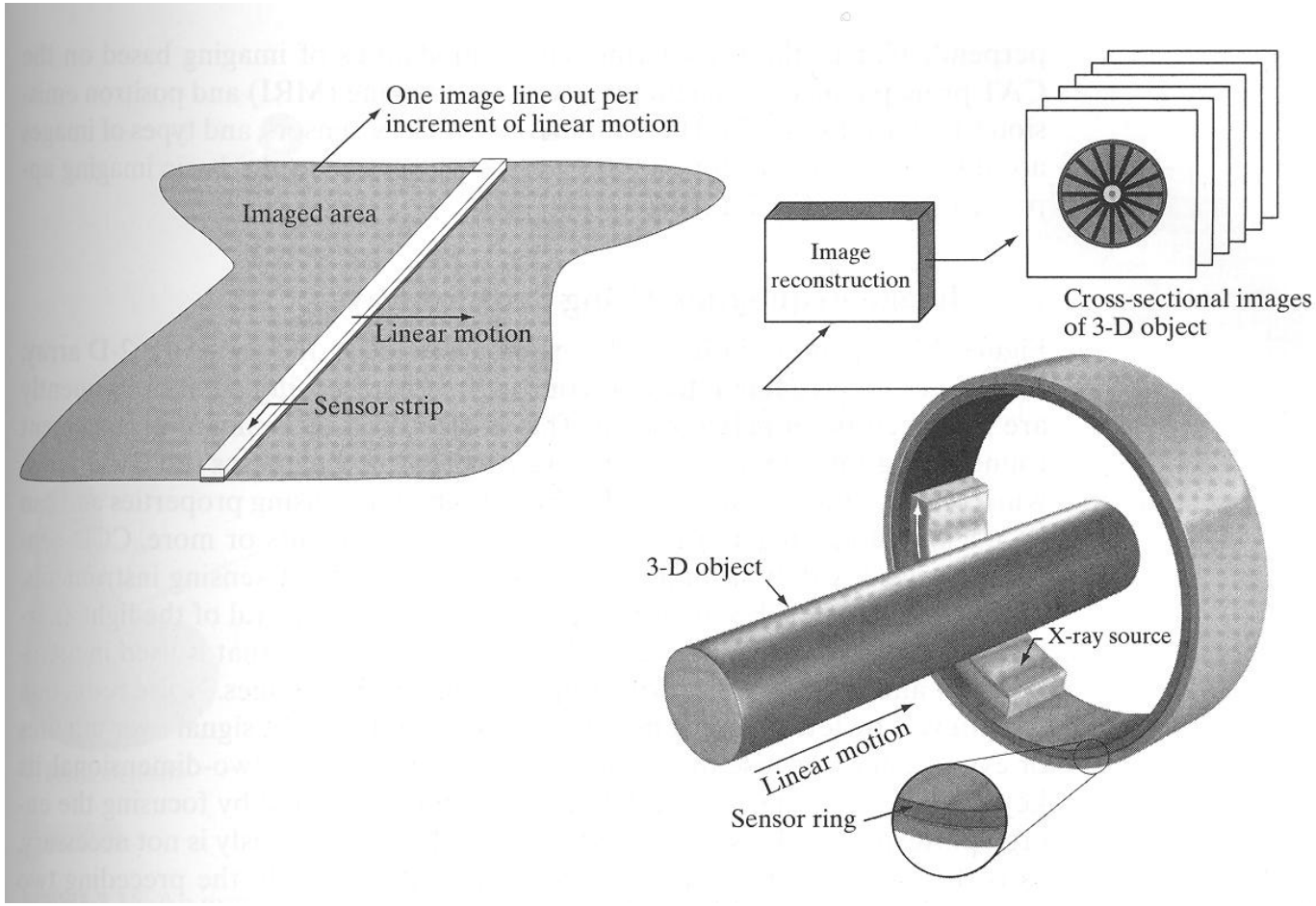


Yhdelläkin anturilla oli mahdollista saada kaksiulotteinen (2D) kuva toimittaessa kuvan 2.10. järjestelyin, jolloin anturi pyyhkäisee ”rivin” kerrallaan vaakasuuntaisena liikkeenä ja rumpu pyörii siirtäen seuraavan rivin kohdalle. Tässä filminegatiivi oli rullalla, jolle saadaan tarkka kuva kuvattavasta kohteesta, koska mekaaninen liike on tarkasti hallittavissa. Tosin menettely oli hidas.

Anturiliuskaa käytetään kaksiulotteisten ja kolmiulotteisten (3D) kuvien kuvaamisessa. Kuva 2.11. esittää periaatteen, jota käytetään useimmissa skannereissa eli kuvanlukijoissa ja toisena tietokonetomografiassa (CT tai CAT, computerized [axial] tomography), magneettiresonanssikuvauksessa (MRI) ja positroniemissiotomografiassa (PET). Tietokonetomografiassa röntgensäde kulkee läpi kohteen ja tulos rekisteröidään renkaassa olevassa anturiliuskassa, joka liikkuu lineaarisesti. Täten saadaan kerroskuvaus aikaiseksi kolmiulotteisena tilavuuskuvana.



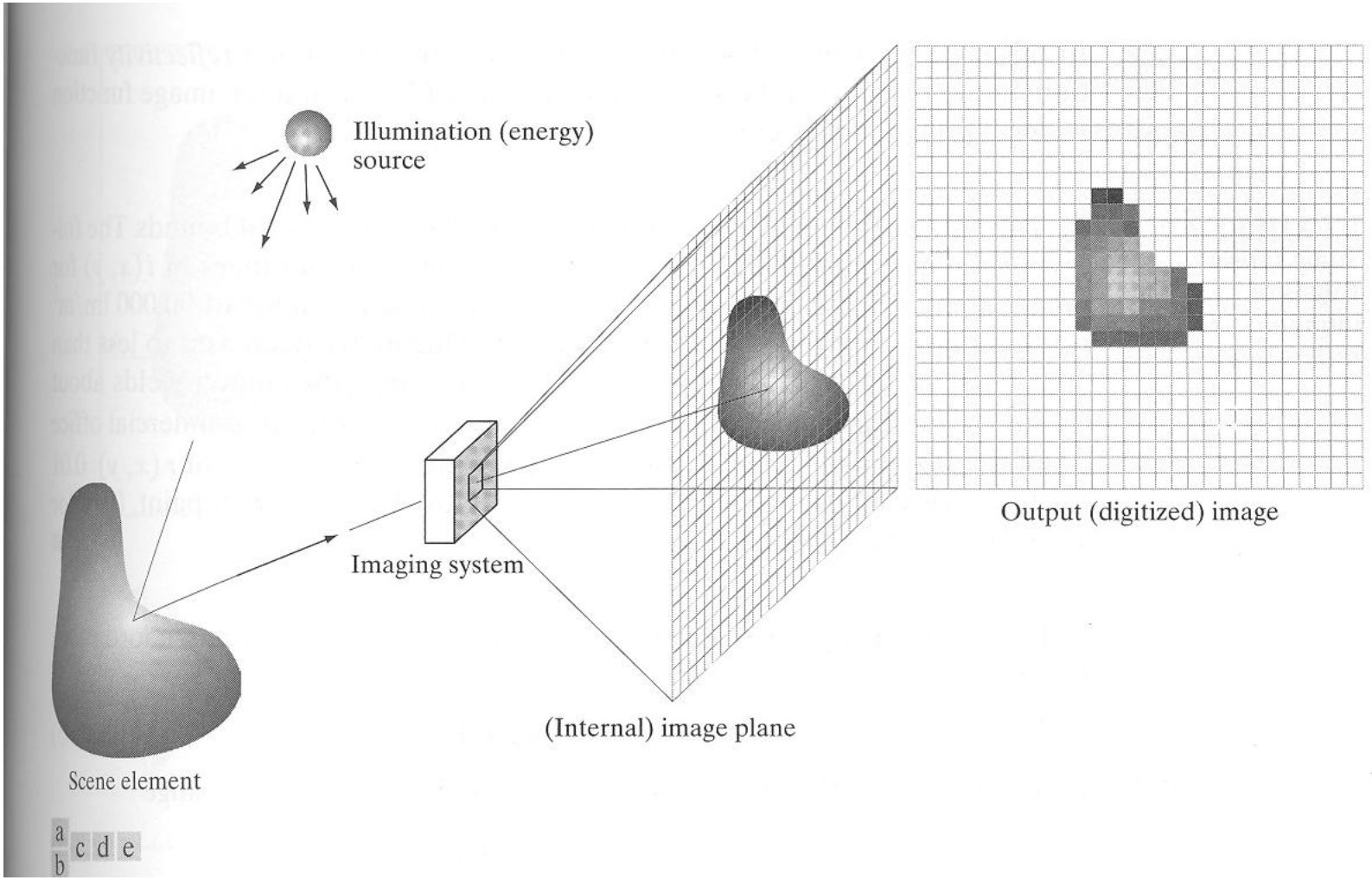
Kuva 2.10. Yhdellä anturilla saadaan kaksiulotteinen kuva filmille.



Kuva 2.11. Kaksi- ja kolmiulotteisen pyyhkäisykuvan luonti.

Kuvan 2.9. anturimatriisia sovelletaan lukuisissa sähkömagneettisissa, joissakin ultraäänilaitteissa ja erityisesti digitaalisissa kameroissa. Kameroiden CCD-kennot (charge-coupled device) voidaan valmistaa tiiviissä muodossa, esim. 4000×4000 elementtiä tai runsaasti enemmän. Matriisijärjestelyn etu liuskaan verrattuna on, ettei tarvita mekaanista liikettä. Tällöin voidaan käyttää pitkää ”valotusaikaa” eli kerätä valoa säteilevän kohteen energiaa suhteellisen pitkään, mikä on tarpeen varsinkin tähtitieteellisissä kuvauksissa. CMOS (Complementary Metal Oxide Semiconductor) on kuitenkin nykyään yleisin kennotyyppi.

Kuva 2.12. havainnollistaa kuvausmenettelyn. Digitaalikamera tai kuvauslaite yleensä rekisteröi valoa tai säteilyä heijastavan näkymän CCD-kennolleen. Jos kyseessä on valoa, järjestelmässä on optinen linssi, joka kerää sitä. Digitaali- ja analogiapiirit pyyhkäisevät kennolta tuloksen, joka digitoidaan kuvauslaitteessa.



Kuva 2.12. Kuvanhankintaprosessi: (a) energia- (valaistus) lähde, (b) näkymäalkio, (c) kuvauslaite ja –järjestelmä, (d) näkymän projektio kuvatasolle ja (e) digitoitu kuva.

Aluksi määriteltiin (kaksiulotteisen) kuvan olevan kaksiulotteinen funktio *f*(*x*,*y*). Amplitudi- eli funktion *f* arvo tasokoordinaateille (*x*,*y*) on positiivinen skalaarisuure, jonka fysikaalisen merkityksen määrää kuvalähde. Kun kuva luodaan fysikaalisella prosessilla, intensiteettiarvot ovat suhteessa säteilyenergiaan, esim. sähkömagneettiset aallot. Näin ollen arvon *f*(*x*,*y*) on oltava ei-negatiivinen ja äärellinen, ts.

$$0 < f(x, y) < \infty.$$

Funktio *f*(*x*,*y*) voidaan jakaa kahteen osaan, (1) valaistuslähteen valo, joka osuu kuvattavaan näkymään (kohteisiin), ja (2) tästä heijastuva valo. Nämä ovat *valaistus i*(*x*,*y*) ja *heijastuvuus r*(*x*,*y*).

Tällöin on

$$f(x,y) = i(x,y)r(x,y),$$

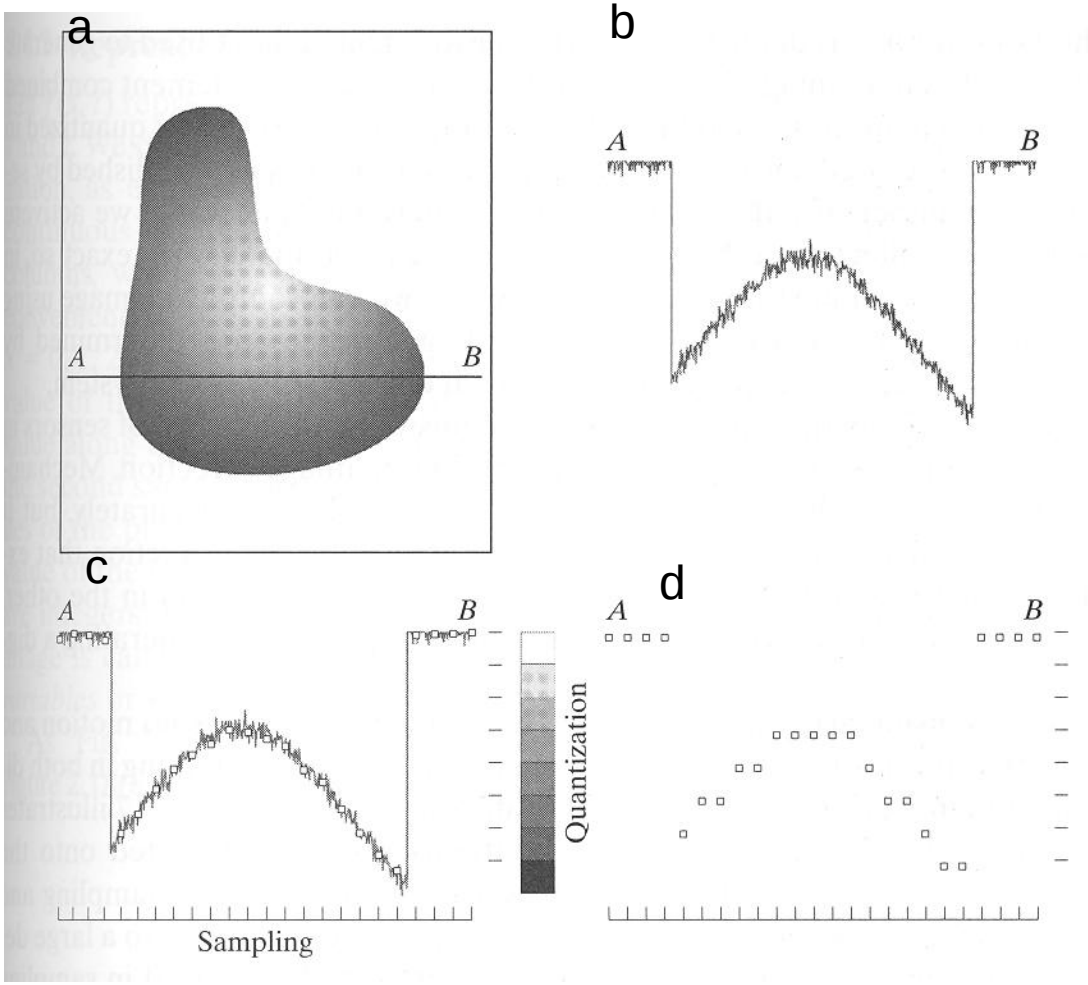
jossa

$$0 < i(x,y) < \infty \text{ ja } 0 < r(x,y) < 1.$$

Heijastuvuus rajoittuu täydellisen absorboinnin 0 ja heijastuksen 1 väliin. Valaistus riippuu valaistuslähteestä ja heijastuvuus kuvattavasta kohteesta. Nämä soveltuvat niin ikään ”läpivalaisuun”, kuten röntgensäteet, vaikka silloin ei voi puhua heijastumisesta. Kirkkaalla säällä aurinko säteilee määrällä 90 000 lm/m^2 ja pilvisellä 10 000 lm/m^2. Tyypillisiä heijastuvuusarvoja ovat 0.01 mustalla sametilla, 0.65 ruostumattomalla teräksellä ja 0.80 tasaisella valkoisella tapetilla ja 0.93 puhtaalla lumella.

2.4. Kuvan näytteistys ja kvantisointi

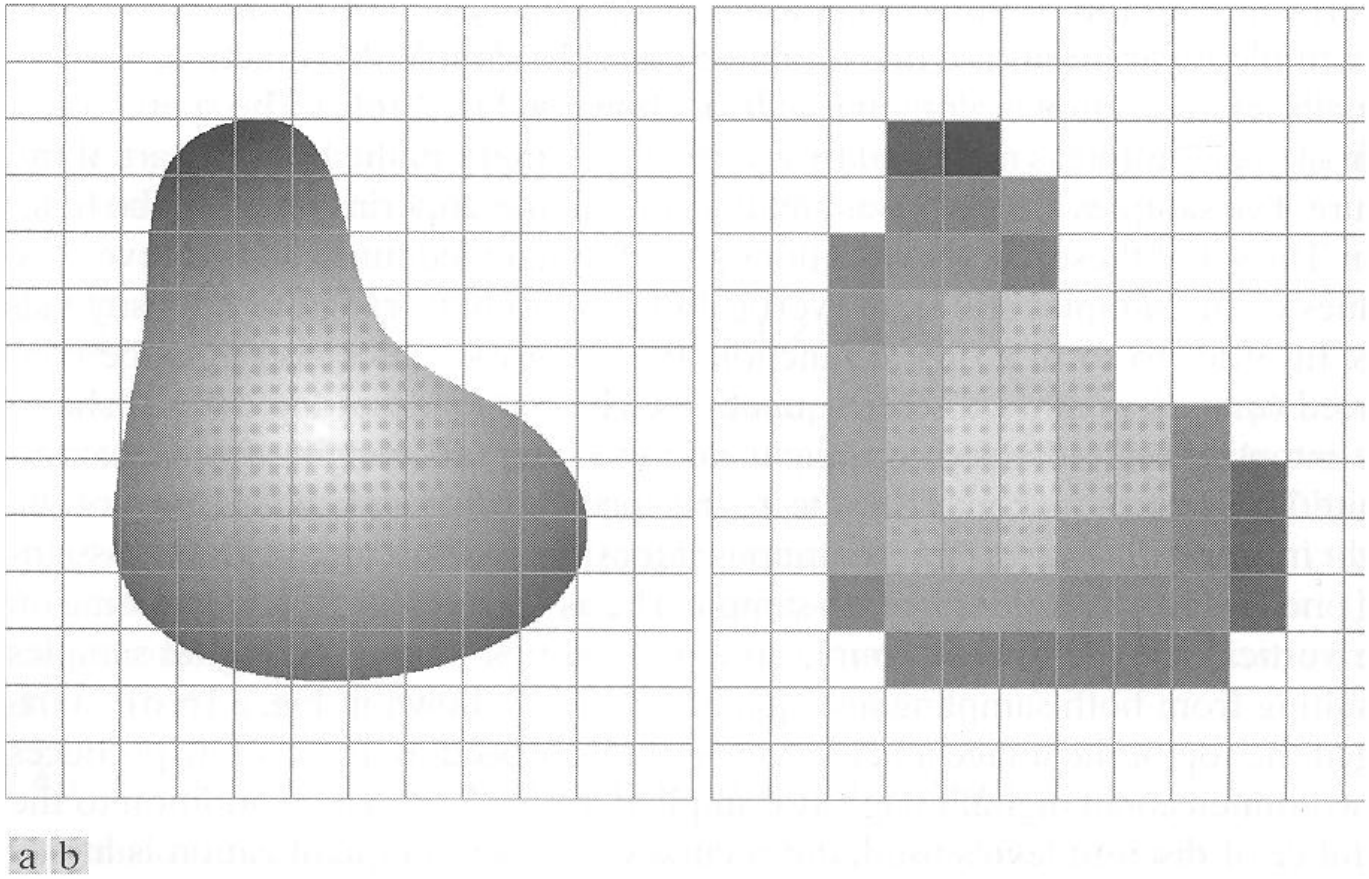
Näytteistyksen ja kvantisoinnin idea esitetään kuvassa 2.13. Vasen yläosa esittää jatkuvan kuvan *f*, joka halutaan muuntaa digitaaliseen muotoon. Kuva voi olla jatkuva *x*- ja *y*-koordinaattien ja myös amplitudin suhteen. Digitalisointia varten jatkuva funktio on näytteistettävä kummankin koordinaatin ja amplitudin suhteen. Koordinaattien digitointi on *näytteistämistä* (sampling) ja amplitudin *kvantisointia* (quantisation). Kuvan 2.13. oikean yläosan yksiulotteinen funktio vastaa jatkuvan kuvan amplitudiarvoja (intensiteetti) kuljettaessa janaa AB pitkin. Mukana oleva satunnaisvaihtelu on kohinaa eli häiriötä. Näytteet otetaan määrävälein janalta AB kuvan 2.13. vasemman alaosan mukaan. Näytteet on esitetty pieninä neliöinä signaalin ”verhona”. Näiden intervallikohdat vastaavat koordinaatteja eli näytteistyspisteitä janalla AB. Lopuksi pitää mitatut jatkuvat arvot diskretisoida eli kvantisoida, so. muuttaa jatkuvat intensiteetit diskreeteiksi, tässä tapauksessa vain kahdeksalle kvantisointitasolle (kuvan oikea alaosa). Näiden määrä muodostaa kuvan amplituditarkkuuden.



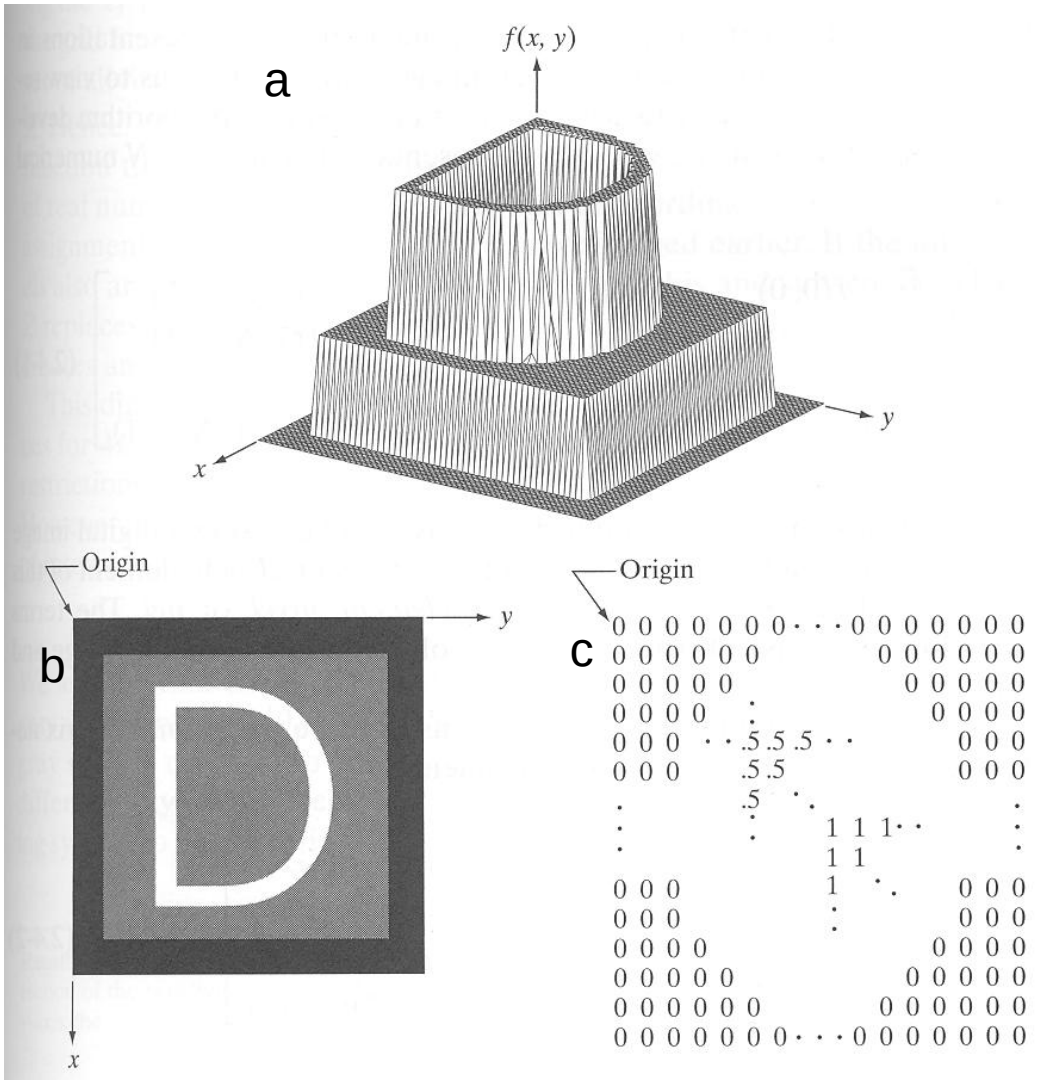
Kuva 2.13. Digitaalisen kuvan generointi: (a) jatkuva kuva, (b) selausjana AB, jolla suoritetaan yhden raidan (c) näytteistys ja (d) saatujen amplitudiarvojen kvantisointi.

Kuva 2.14. esittää koko kuvallista tietoa. Kuvan laatu, joka tässä on havainnollistuksen vuoksi sangen harva, määräytyy olennaisin osin käytettyjen näytteiden ja intensiteettitasojen määrän, ts. tarkkuuden eli resoluution mukaan.

Merkitään näytteistettyä ja kvantisoitua kuvaa $f(x,y)$, jossa $x=0,\ldots,M-1$ ja $y=0,\ldots,N-1$. Kuvassa 2.15. on kolme vaihtoehtoa esittää $f(x,y)$. Kuvan yläosassa on kolmiulotteinen (todellisuudessa sen projektio) vaihtoehto, alhaalla vasemmalla miltä se katsojasta näyttäisi (mahdollisesti ilman värejä) ja oikealla alhaalla ”numeerinen” esitys.



Kuva 2.14. (a) Jatkuva kuva projisoituna matriisille sekä (b) näytteistyksen ja kvantisoinnin tulos.



Kuva 2.15. (a) Kuvan pintakuvaus, (b) kuva visuaalisena ja (c) kvantisoituina lukuarvoina.

Kuvan 2.15 vasen alaosa on luonnollisesti visuaalisesti kiinnostava versio ja mitä näytöllä näytetään, mutta kaksi muuta kertovat, mitä se datana sisältää. Tässä oli vain kolme kvantisointiarvoa käytössä vastaten harmaasävyjä: musta, harmaa ja valkoinen. Jos kuvassa on 4000×4000 alkiota, pikseliä, se sisältää kaikkiaan 16 miljoonaa kvantisoitua arvoa. Kuvan oikeassa alalaidassa nähdään kirjoitetun vain osa kaikista luvuista. Ideana on, paitsi havainnollistavuus, myös tuoda mieleen, että kuvien algoritmisessa käsittelyssä pyritään samalla tavalla eliminoimaan osa datasta prosessin nopeuttamiseksi ja ylipäänsä enemmän tai vähemmän ”turhana” laskentana.

Kuva *f*(*x*,*y*) ymmärretään nyt pikselimatriisina, esim. esityksenä (indeksit voisivat olla myös alaindekseinä).

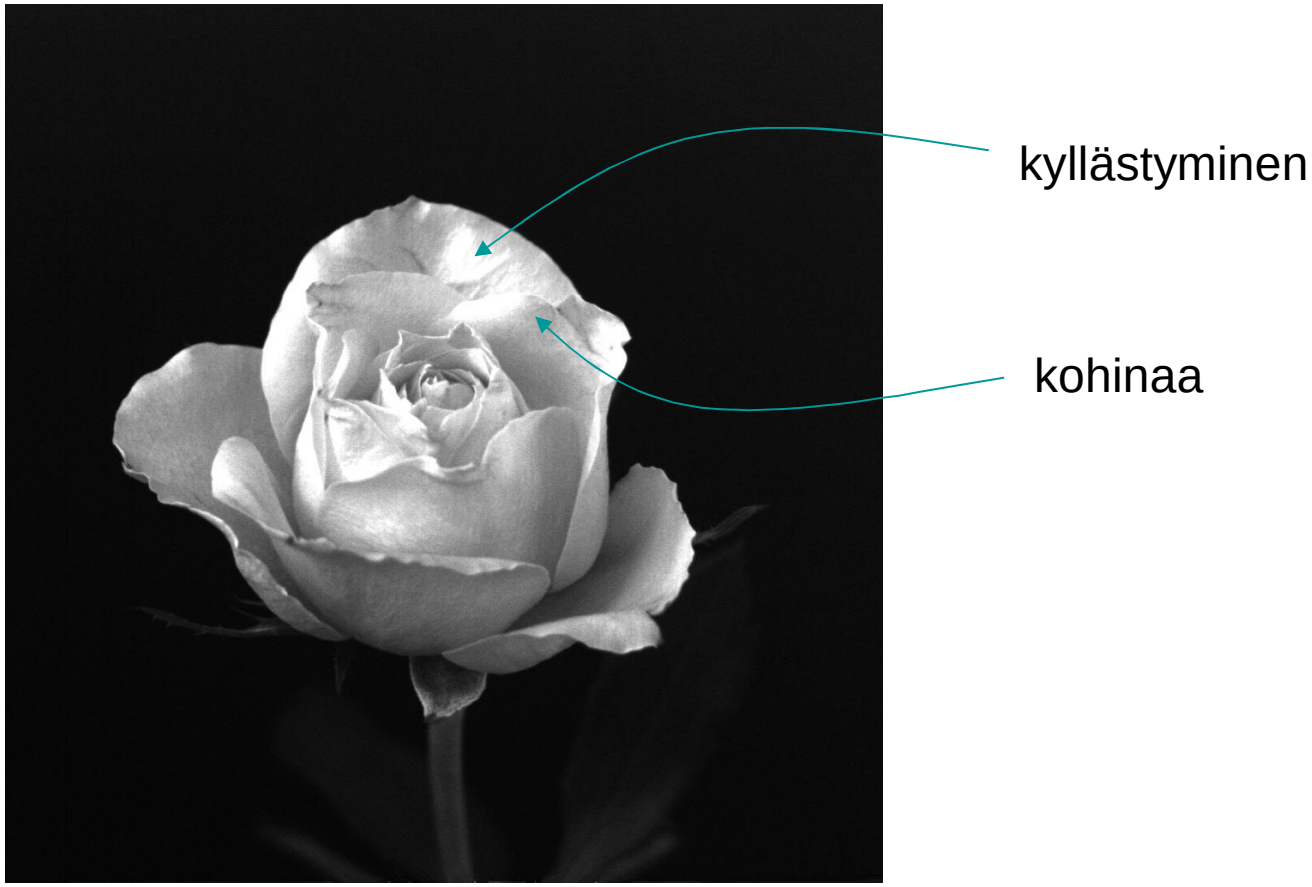
$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix}$$

Ajateltaessa tätä kuvana huomataan origon olevan vasemmassa yläkulmassa ja akselien kulkevan vasemmalta oikealle ja alas. Monet näytöt, esim. TV, pyyhkäisevät raita kerrallaan alkaen vasemmalta ylhäältä. Kyseessä on oikeakätinen karteesinen koordinaatisto.

Kuvan digitointia varten on päätettävä määrät *M* ja *N* sekä kvantisointi- eli intensiteettitasojen lukumäärä *L*. Muistin käytön kannalta on tavallisesti mielekästä määrittää *L* 2:n potensseina, ts.

$$L = 2^k.$$

Oletetaan diskreettien tasojen olevan tasavälisiä ja kokonaislukuja lukuväliltä [0,*L*-1]. Tätä kutsutaan mitattavan suureen, tässä intensiteetti, *dynaamiseksi alueeksi.* Maksimi eli yläraja voidaan ymmärtää kyllästymispisteenä ja alaraja (kvantisointi)kohinan rajana (kuva 2.16). *Kontrasti* määritellään kuvassa esiintyvän suurimman ja pienimmän intensiteettiarvon erotuksena. Jos dynaaminen alue on laaja, kuvassa voi olla suuria kontrasteja. Jos se on pieni, tuloksena on haalea kuva.



Kuva 2.16. Kyllästyminen (saturaatio) tarkoittaa, että dynaamisen alueen yläraja on ylitetty, jolloin kaikki sen yläpuolelle menneet arvot leikkautuvat tähän maksimiin, ts. saavat kaikki saman maksimiarvon. Kohina on (yleensä) pieniamplitudista satunnaisilmiötä, joka ”sotkee” kuvaa varsinkin dynaamisen alueen alarajalla, ts. ”maskaa” tätä. Kohina muodostaa tässä rakeisia tekstuureita, erikoisesti tummissa kohdissa

Eo. merkinnöin saadaan *b*, kuvan tallentamiseen tarvittava bittien lukumäärä

$$b = M \times N \times k.$$

Kun on *M*=*N*, edellinen on yhtä kuin

$$b = N^2 k.$$

Taulukko 2.1. antaa kuvien koot bitteinä eri *N*:n, *k*:n ja *L*:n arvoilla. Tavun ollessa 8 bittiä kuvien vaatimat muistitilamäärät kasvavat nopeasti näiden ollessa suurempia kuin 1024×1024 pikseliä.

Taulukko 2.1.

Number of storage bits for various values of N and k. L is the number of intensity levels.

N/k	1 ($L = 2$)	2 ($L = 4$)	3 ($L = 8$)	4 ($L = 16$)	5 ($L = 32$)	6 ($L = 64$)	7 ($L = 128$)	8 ($L = 256$)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

Spatiaalinen eli tilaa koskevaa resoluutio on kuvan pienimmän erotettavissa olevan yksityiskohdan yksikkö. *Spatiaalinen resoluutio* esitetään yleensä suureena pistettä tai pikseliä mittayksikköä kohti (dots per inch, dpi), jossa käytetään USA:n tapaan tuumaa. Esim. sanomalehdet on painettu resoluutiolla 75 dpi, aikakauslehdet 133 dpi ja korkealaatuiset kirjat 2400 dpi.

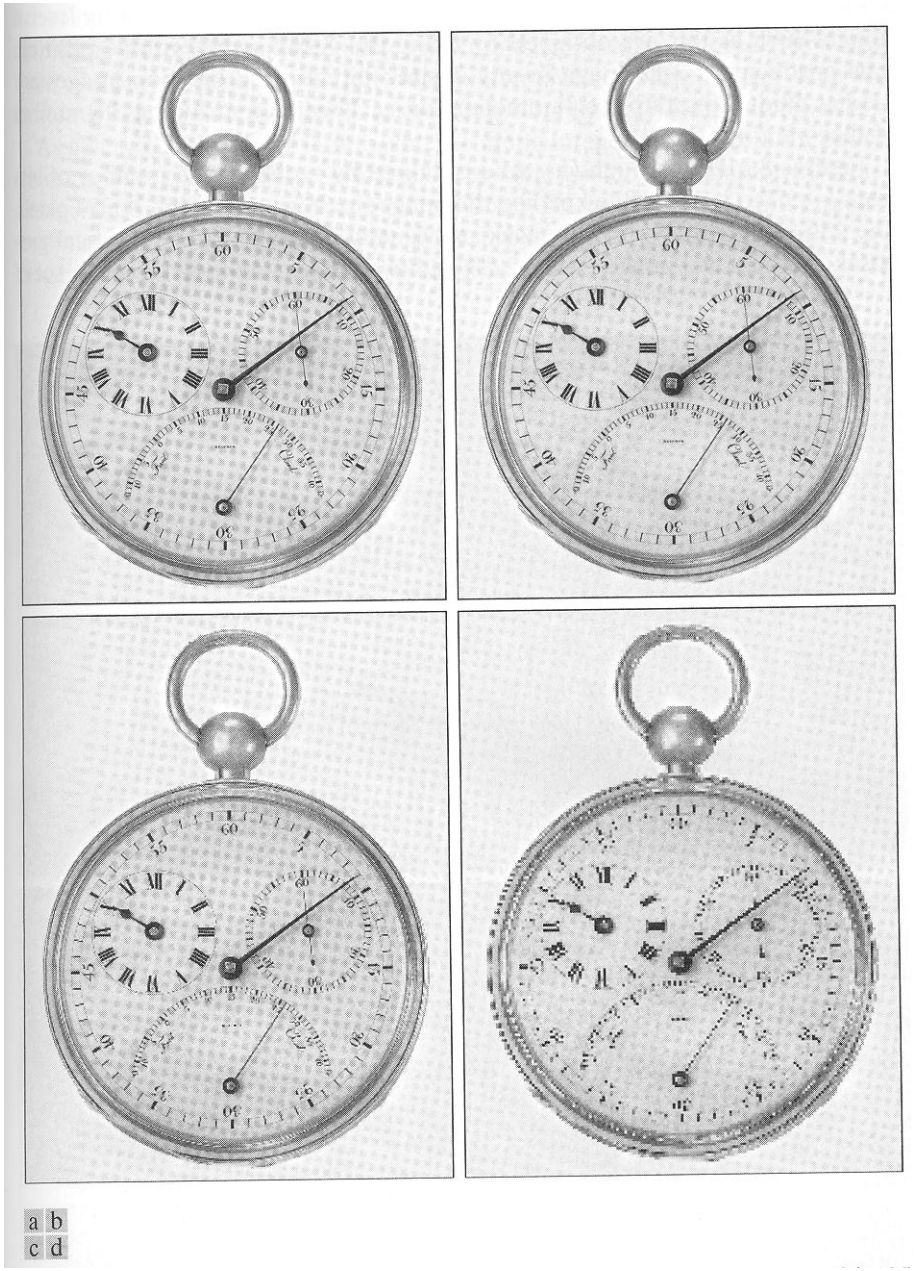
Olennaista on suhteuttaa spatiaalinen resoluutio käytettävien (absoluuttisten) mittayksiköiden kanssa. Sanottaessa kuvan resoluution olevan 1024×1024 pikseliä tämä ei vielä kerro kaikkea, jos ei tiedetä spatiaalista resoluutiota. Pikselimäärillä voidaan vain karkeasti verrata eri laitteiden tarkkuutta, kuten kameroissa, joissa nämä ovat 20 megapikseliä tai 8 megapikseliä olettaen samalla näissä olevan samanlainen linssi ja kuvan otettavan samalta etäisyydeltä.

Intensiteettiresoluutio viittaa vastaavasti intensiteettitasojen pienimpään eroon. Tämä on yleensä 2:n potenssi tullen muistiin tallentamisen kannalta sopivista bittien lukumääristä. Tavallisin on 8 bittiä (1 tavu), joskus 16 bittiä tai jopa 32 bittiä. Harvinaisina myös 10 ja 12 bittiä on käytetty. Tällöin 8 bittiä vastaa luonnollisesti 2^8=256 tasoa. Kun kuvan kohina ja kyllästyminen vaikuttavat merkittävästi kuvan laatuun, resoluutio on tässäkin ainoastaan lähtökohta lopputuloksen kannalta.

Kuva 2.17. esittää spatiaalisen resoluution vähentämisen vaikutuksia. Alkuperäistä 3692×2812 alempi resoluutio antaa luonnollisesti pienempiä kuvia, esim. kuvan 72 dpi koko 213×162. Vertailun vuoksi pienemmät kuvat on suurennettu alkuperäisen kokoiseksi.

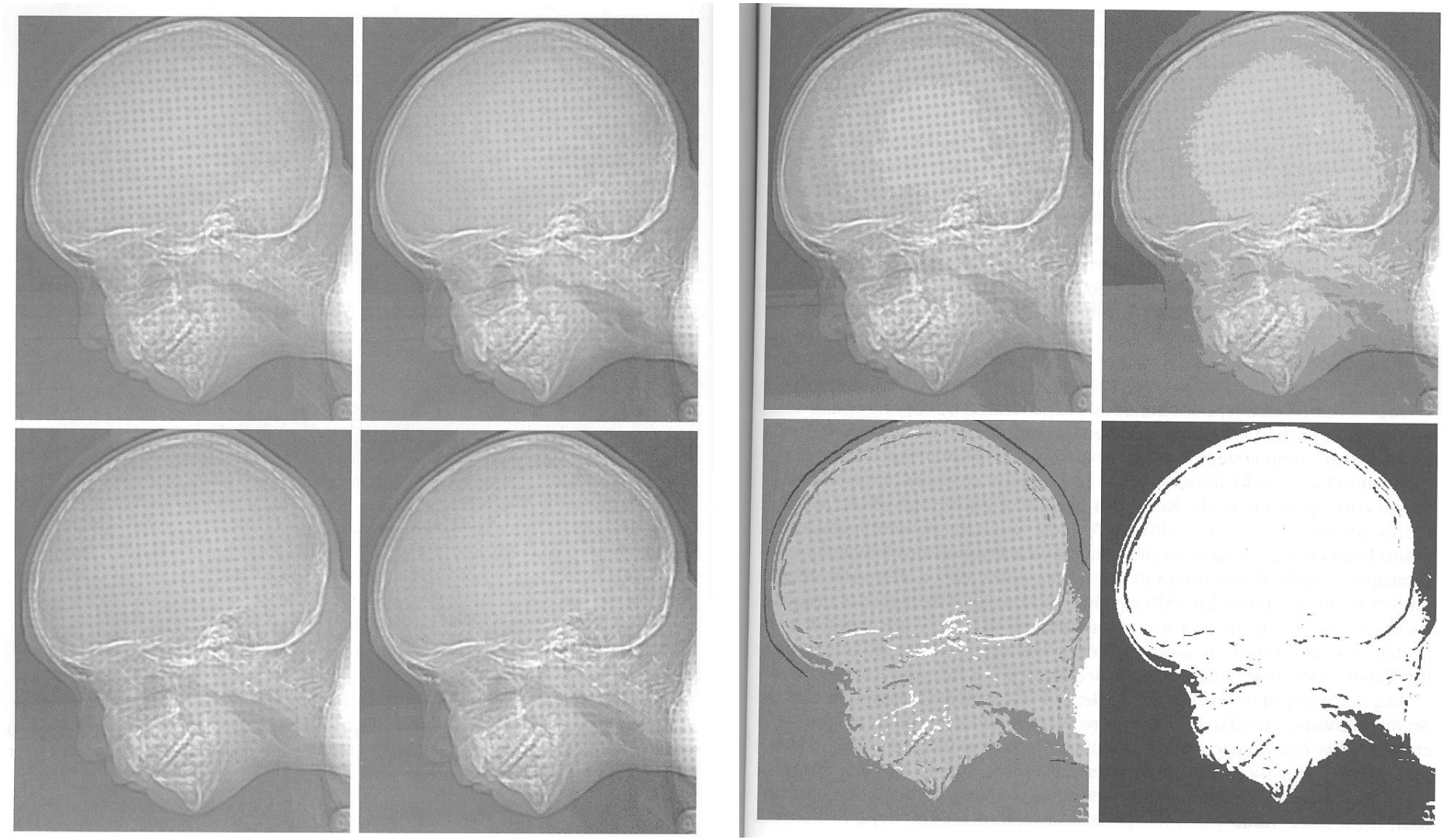
Kuva 2.17. Spatiaalisen resoluution vähentämisen vaikutuksia: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi ja (d) 72 dpi.

300 dpi on melko hyvä, ja se onkin alin kirjojen kuvissa käytetty. Tapauksessa 150 dpi kaarevat reunat ja mm. kronometrin oikeanpuoleisen pienen kellotaulun viisari ovat huonoja. 72 dpi on varsin kehno. Tällaisia puutteita on kuitenkin mahdollista ”tempuilla” korjata esim. lokaalisti vaihtelemalla pikselin kokoa.



Tarkastellaan seuraavaksi intensiteetin vähentämisen vaikutuksia. Kuvassa 2.18. on vähennetty intensiteettiä arvosta 256 (*k*=8) arvoon 2 eli mustavalkokuvaan. Alkuperäinen on CT-kuvan kaksiulotteinen projektiokuva kokoa 452×374, jossa on 256 intensiteettitasoa. Kuvasarjassa vähennettiin arvoa *k*=8 arvoon 1. Sen sijaan pikselimäärä pidettiin muuttumattomana.

Intensiteetit 256, 128 ja 64 ovat visuaalisesti samanlaisia. Tapauksesta 32 voi jo havaita hyvin vähäisiä harjannemaisia rakenteita kohdissa, joissa on vakio- tai lähes vakiointensiteettiä (tässä kuvan skannauksen takia kuva ei ole aivan ”aito”, esim. rakeisuus on tullut skannauksesta). Näitä kutsutaan *vääriksi ääriviivoiksi.* Ilmiö esiintyy selvästi edelleen alemmissa intensiteeteissä. Tässä yhteydessä 64 olisi ollut riittävä intensiteettitasojen määrä.



ab ef
cd gh

Kuva 2.18. Intensiteettitasojen määrä: (a) 256, (b) 128, (c) 64, (d) 32, (e) 16, (f) 8, (g) 4 ja (h) 2.

Edellä pohdittiin suureiden *N* ja *k* vaikutusta erikseen. Todellisuudessa näillä on yhteisvaikutus, jota on kuitenkin vaikeaa mitata. Erilaisia subjektiivisia (katsojan) arviointeja on tehty vaihteleville kuville, kuten muutenkin kuvankäsittelytutkimuksessa runsaasti käytetyille testikuville, jotka esitetään kuvassa 2.19. Nähdään, miten kuvan sisällöllä on tietysti merkitystä. Esim. kuvassa, jossa on vähän yksityiskohtia ja reunoja sekä paljon tasaista yhtenäistä pintaa, pärjätään ehkä vaatimattomillakin resoluutiolla. Toisaalta paljon yksityiskohtia sisältävä kuva on vaativa.

(a)

(b)

(c)

Kuva 2.19. (a) ”Lena”, vähän yksityiskohtia, (b) ”cameraman”, keskimääräinen kuva yksityiskohtiensa lukumäärän suhteen ja (c) paljon yksityiskohtia [MIT].

Kuvan interpolointi on tärkeä perusmenetelmä, jota käytetään mm. kuvan zoomauksessa, kutistamisessa, rotaatiossa ja geometrisissä korjauksissa. Se on pohjimmiltaan *uudelleennäytteistys*menetelmä (resampling).

Interpoloinnissa käytetään tunnettuja arvoja tai pisteitä tuottamaan arvoja ”tuntemattomissa” paikoissa, niiden välissä. Esim. kun alkuperäinen kuva on kokoa 500×500 ja halutaan zoomata se kokoon 750×750, luodaan aluksi ristikko kokoa 750×750, joka kutistetaan samaan pinta-alaan kuin alkuperäinen. Nyt kukin pikseli peittää luonnollisesti pienemmän osan kuin alkuperäisessä. Uusien pikselien intensiteettiarvot saadaan vertaamalla niitä alkuperäisiin pikseleihin. Kopioidaan lähimmästä alkuperäisestä intensiteetti kuhunkin uuteen pikseliin. Lopuksi suurennetaan kuva zoomattuun kokoonsa. Kyseessä on *lähimmän naapurin interpolointi.*

Lähimmän naapurin interpolointi on yksinkertainen toteuttaa, mutta se on taipuvainen tuottamaan *artefaktoja*, kuten suorien vääristymiä. Näin ollen käytännössä sovelletaan muita, esim. *bilineaaria interpolointia*, jossa estimoidaan neljän lähimmän naapurin intensiteettiarvolla annetun paikan (*x*,*y*) intensiteetti. Se voidaan ajatella esim. ristikon koordinaatteina. Olkoon *v*(*x*,*y*) sen intensiteettiarvo. Tämä saadaan kaavalla

$$v(x, y) = ax + by + cxy + d, \qquad (3)$$

jossa neljä kerrointa lasketaan tuntemattomina neljän lähimmän naapurin antamasta neljästä yhtälöstä tunnettaessa näiden intensiteetit. (Vakiintuneen nimen osa bilineaari on kehnosti valittu, kun kaavassa on mukana myös muuttujien *x* ja *y* sekatermi, so. matemaattisesti epälineaari termi.)

Vaikeustasoltaan seuraava menetelmä on *kaksoiskuutiollinen interpolointi* (*bicubic interpolation*). Tässä on mukana 16 lähintä naapuripistettä. Pisteen (*x*,*y*) intensiteettiarvo lasketaan kaavalla

$$v(x,y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i y^j , \qquad (4)$$

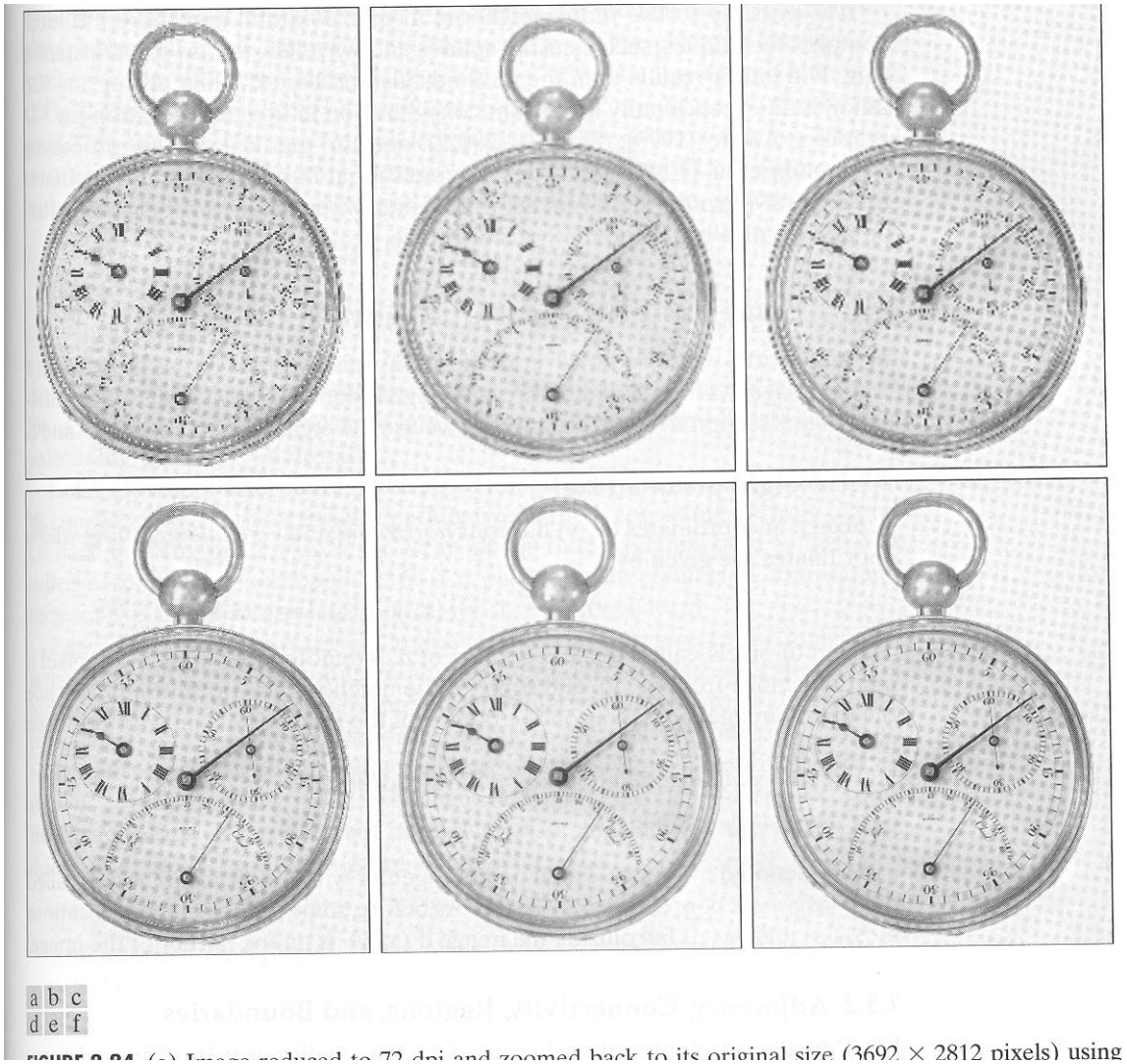
jossa 16 kerrointa määrätään tuntemattomina 16 naapurin avulla. Edeltävän kaavan (3) termit ovat osa (4):stä, joka kerää näin informaatiota laajemmalta alueelta ja säilyttää monesti edellistä paremmin yksityiskohdat. Kaavaa (4) käytetään Adobe Photoshopissa ja Corel Photopaintissa.

Palataan kuvan 2.17 esimerkkiin, jonka osan (d) luomiseksi, sekä kutistus että zoomaus, oli käytetty lähimmän naapurin interpolointia. Tämä on nyt myös kuvassa 2.20, osana (a), joka on kehno. Kuvan osat (b) ja (c) on saatu käyttäen bilineaaria ja kaksoiskuutiollista interpolointia, jolloin parannus on huomattava. Jälkimmäinen on lisäksi edellistä hiukan terävämpi.

Kuvan 2.20. osa (d) on sama kuin kuvan 2.17 (c), jolle on tehty vastaavasti interpoloinnit kuin yllä kuvan 2.20. osissa (e) ja (f). Huolimatta resoluution vähenemisestä alkuperäisestä 1250 niinkin alas kuin 150 osien (e) ja (f) kuvat ovat melko hyviä, mikä osoittaa interpoloinnin tehokkuutta, varsinkin kaavalla (4).

Näitä monimutkaisempia menetelmiä ovat splinejä tai aallokkeita hyödyntävät, joita sovelletaan kuitenkin vain erikoistapauksissa, mm. joskus lääketieteellisten kuvien yhteydessä.

Kuva 2.20. (a) Kuva on supistettu tarkkuuteen 72 dpi ja zoomattu takaisin alkuperäiseen kokoonsa soveltaen lähimmän naapurin interpolointia, (b) kuva kutistettuna ja zoomattuna soveltaen bilineearia interpolointia, (c) kuten edellinen, mutta kaksoiskuutiollisen interpoloinnin avulla. (d)-(f) Edellisiä vastaava kuvajono, mutta alussa oli supistettu tarkkuuteen 150 dpi.



a b c
d e f

FIGURE 2.24 (a) Image reduced to 72 dpi and zoomed back to its original size (3692 × 2812 pixels) using

2.5. Pikselien välisiä suhteita

Kuvataan kuvaa jälleen funktiona *f*(*x*,*y*). Pikselillä *p* koordinaateilla (*x*,*y*) on neljä horisontaalista ja vertikaalista naapuria, joiden koordinaatit ovat seuraavat

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1).$$

Tätä pikselijoukkoa kutsutaan pikselin *p 4-naapureiksi*, $N_4(p)$. Jokainen näistä neljästä on yksikön päässä *p*:stä, mutta reunalla niitä voi mennä jo kuvan ulkopuolelle. Pikselin neljä diagonaalista naapuria, $N_D(p)$, ovat seuraavat.

$$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1).$$

Nämä kaikki yhdessä muodostavat *p*:n *8-naapurit*, $N_8(p)$, joista jotkut voivat olla kuvan reunalla.

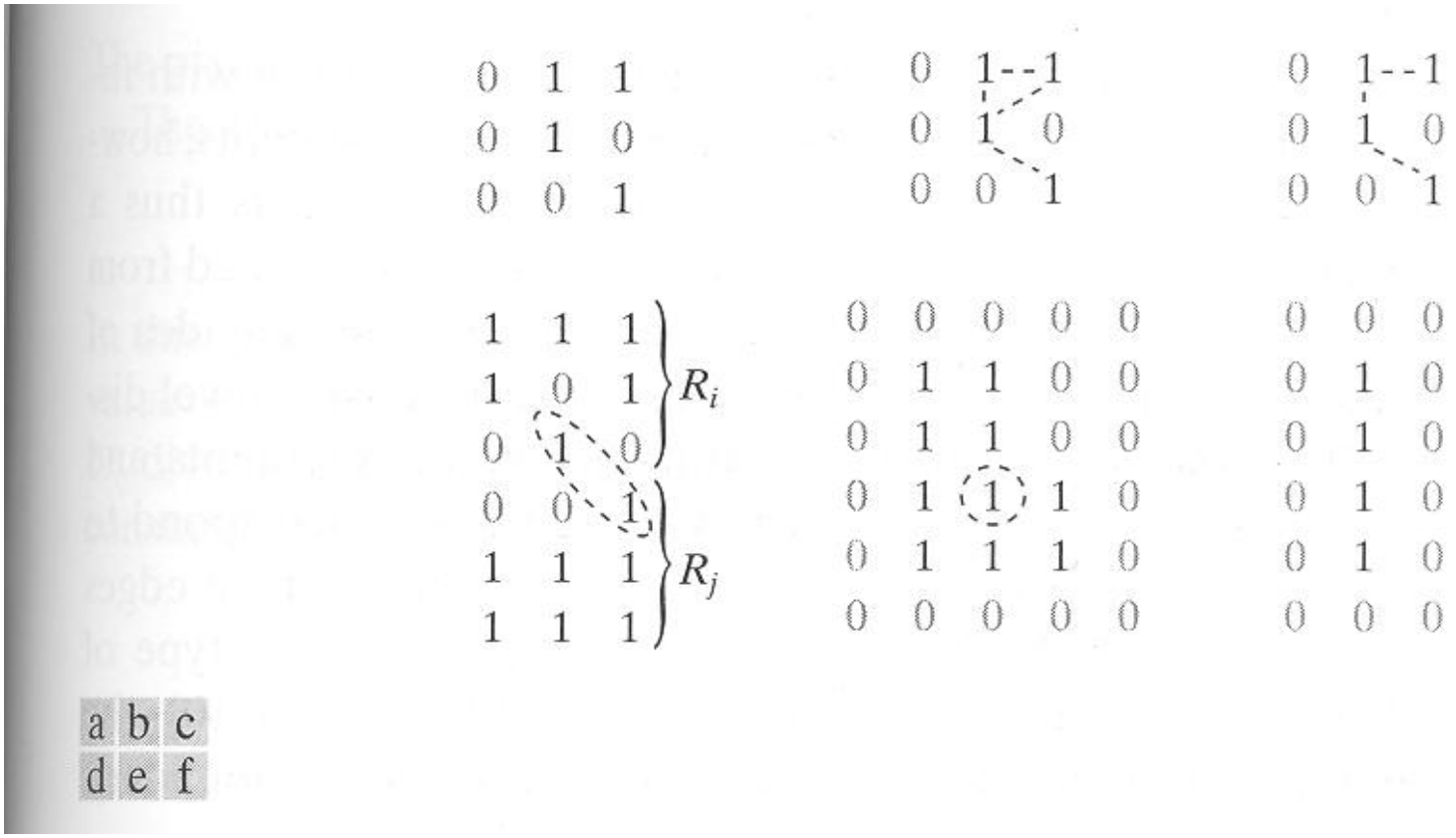
Olkoon V intensiteettiarvojen joukko, jota käytetään määrittelemään vierekkäisyyttä. Esim. harmaasävyillä {0,1,…,255} V on jokin osajoukko näistä ja binäärillä V={1}. Tarkastellaan kolmea eri vierekkäisyyttä:

(1) *4-vierekkäisyys*: Pikselit p ja q ovat *4-vierekkäisiä*, jos niiden intensiteetit kuuluvat osajoukkoon V ja q on $N_4(p)$:ssä.

(2) *8-vierekkäisyys*: Pikselit p ja q ovat *8-vierekkäisiä*, jos niiden intensiteetit kuuluvat osajoukkoon V ja q on $N_8(p)$:ssä.

(3) *sekavierekkäisyys*: Pikselit p ja q ovat *sekavierekkäisiä*, jos niiden intensiteetit kuuluvat osajoukkoon V ja joko (i) q on $N_4(p)$:ssä tai (ii) q on $N_D(p)$:ssä ja joukolla $N_4(p) \cap N_4(q)$ ei ole yhtään pikseliä, jonka arvo olisi osajoukossa V.

Sekavierekkäisyys on muunnos 8-vierekkäisyydestä. Sen tarkoitus on hävittää ei-yksikäsitteiset tilanteet, joita 8-vierekkäisyys saattaa aiheuttaa. Tästä on esimerkkinä kuva 2.21.



Kuva 2.21. (a) Pikselijärjestys, (b) pikselit, jotka ovat 8-vierekkäisiä (merkitty katkoviivoilla; huomaa ei-yksikäsitteisyys), (c) sekavierekkäiset, (d) kaksi 1-aluetta, jotka ovat 8-vierekkäisiä, (e) ympyröity piste on osa 1-arvoisten pikselien rajaa vain, jos 8-vierekkäisyyttä käytetään 1-alueen ja taustan välillä ja (f) 1-arvoisten alueen sisäraja ei muodosta suljettua *polkua*, kuten ulkoraja tekee.

Polku koordinaattien (x,y) pikselistä p koordinaattien (s,t) pikseliin q on pikselien jono

$$(x_0,y_0), (x_1,y_1), \ldots,(x_n,y_n),$$

jossa $(x_0,y_0)= (x,y)$, $(x_n,y_n)=(s,t)$ sekä pikselit (x_i,y_i) ja (x_{i-1},y_{i-1}) ovat vierekkäisiä, kun $1 \leq i \leq n$. Tällöin n on polun pituus. Jos $(x_0,y_0)= (x_n,y_n)$, se on *suljettu polku*. Voidaan määritellä 4-, 8- tai sekapolku riippuen vierekkäisyyden tyypistä.

Olkoon S kuvan pikselien osajoukko. Pikselit p ja q ovat *yhdistettyjä*, jos S:ssä on polku niiden välillä. Jokaiselle pikselille p joukossa S niiden pikselien joukkoa, jotka on yhdistetty siihen S:ssä, kutsutaan *yhdistetyksi komponentiksi*. Jos S:ssä on vain yksi sellainen, kyseessä on *yhdistetty joukko.*

Olkoon R kuvan pikselin osajoukko. R on *alue*, jos se on yhdistetty joukko. Kahta aluetta R_i ja R_j sanotaan vierekkäisiksi, jos niiden unioni muodostaa yhdistetyn joukon. Alueet, jotka eivät ole vierekkäisiä, ovat *erillisiä*.

Olkoot R_k, k=1,2,…,K, erillisiä alueita, joista mikään ei kosketa kuvan rajaa (yleisyyttä rikkomatta näin voidaan tehdä, sillä tarvittaessa voitaisiin asettaa yhden pikselin vahvuinen ”vaippa” reunustamaan kuvaa). Olkoon R_u näiden alueiden unioni ja $(R_u)^c$ tämän komplementti. Tällöin R_u on kuvan *edusta* ja $(R_u)^c$ *tausta.*

Raja tai *ääriviiva* muodostuu alueen niistä pisteistä, jotka ovat vierekkäisiä komplementtialueen pisteiden kanssa. Tällöin pitää jälleen määrätä sovellettava yhdistävyys eo. vierekkäisyys-käsitteiden avulla. Käytetään yleensä 8-yhdistävyyttä.

Sisäraja ja *ulkoraja* mainittiin jo kuvan 2.21. yhteydessä. Ulkoraja erottaa sisärajan taustasta. Erolla on merkitys rajaa seuraavien algoritmien käytössä. Tavallisesti käytetään ulkorajaa, jotta saadaan suljettu polku.

Jos alue *R* sattuu olemaan yhtä kuin koko kuva (pikselijoukon nelikulmio), sen raja määritetään kuvan vasemman- ja oikeanpuolisen sarakkeen sekä ylimmän ja alimman rivin mukaan. Normaalisti rajalla olevat pikselit luetaan kuvaan mukaan.

Raja muodostaa suljetun polun ollen globaali luonteeltaan. Sitä vastoin *reuna* (edge) on lokaali käsite vastaten intensiteettiarvojen mitan epäjatkuvuuskohtaa. Tässä pikselien derivaatta-arvot ylittävät jonkin kynnysarvon. Reunapisteitä voidaan yhdistää *reunasegmenteiksi*, jotka joskus saattavat vastata rajaa. Binäärikuvien kohdalla poikkeuksellisesti reunat vastaavat rajoja. Riippuen yhdistettävyyden tyypistä ja reunaoperaattoreista binäärialueesta irrotettu raja voi olla sama kuin alueen raja.

D on *etäisyysfunktio* tai *metriikka* koordinaattien (*x*,*y*), (*s*,*t*) ja (v,w) pikseleille *p*, *q* ja *z* täyttäessään seuraavat ehdot.

(1) $D(p,q) \geq 0$ ja erityisesti $D(p,q)=0$, jos ja vain jos $p=q$ (refleksisiivisyys)

(2) $D(p,q)= D(q,p)$ (symmetria)

(3) $D(p,z) \leq D(p,q)+ D(q,z)$ (kolmioepäyhtälö)

Euklidinen etäisyys määritellään pikselien välillä seuraavasti.

$$D_e(p,q) = \left[(x-s)^2 + (y-t)^2\right]^{\frac{1}{2}}$$

Tällöin pikselit, joiden etäisyys keskipisteestä (*x*,*y*) on enintään annettu *r*, kuuluvat *r*-säteiseen ympyrään tai (ympyrä)levyyn.

D_4-etäisyys (Manhattan- tai korttelietäisyys) määritellään pikselien p ja q välillä. $D_4(p,q) = |x-s| + |y-t|$

Tällöin pikselit, jotka ovat enintään jollakin D_4-etäisyydellä r keskipisteestä (x,y), muodostavat vinoneliön.

```
          2
      2   1   2
  2   1   0   1   2
      2   1   2
          2
```

Pikselit, joilla D_4=1, ovat pikselin (x,y) 4-naapureita.

D_8-etäisyys (šakkilauta-etäisyys) pikselien *p* ja *q* välillä määritellään seuraavalla kaavalla.

$$D_8(p,q) = \max(|x-s|, |y-t|)$$

Tässä etäisyyden *r* keskipisteestä (*x*,*y*) määräämän alueen pikselit muodostavat neliön. Esim. pikselit D_8-etäisyydellä≤2 keskipisteestä muodostavat oheiset rajat.

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Pikselit etäisyydellä D_8=1 ovat keskipisteen 8-naapureita.

Sovellettaessa sekavierekkäisyyttä kahden vierekkäisen pikselin etäisyys riippuu pikselien arvosta polulla ja näiden naapurien arvosta, jolloin tilanne voi vaihdella sen suhteen, mitkä pikselit tulevat mukaan.

2.6. Kuvanprosessoinnin matemaattiset perusvälineet

Taulukko-operaatio toteutetaan pikseleittäin. Selvyyden vuoksi tässä käytetään sanaa taulukko matriisin sijasta. Esim. taulukoille

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ja} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

näiden tulo on nyt

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}.$$

Toisaalta *matriisitulo* on

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

Luentoesityksessä oletetaan käytettävän kautta linjan taulukko-operaatioita, ellei toisin mainita. Esim. korotettaessa taulukko neliöön se tehdään tavallisesti alkioittain nämä neliöimällä.

Lineaaristen ja epälineaaristen menetelmien ero on olennainen. Olkoon *H* jokin kuvalle *f*(*x*,*y*) soveltuva operaattori, joka tuottaa tuloksen *g*(*x*,*y*)

$$H\left[f\left(x,y\right)\right]=g\left(x,y\right).$$

H on lineaarinen, jos

$$\begin{aligned}H\left[a_i f_i(x,y)+a_j f_j(x,y)\right]&=a_i H\left[f_i(x,y)\right]+a_j H\left[f_j(x,y)\right]\\&=a_i g_i\left(x,y\right)+a_j g_j\left(x,y\right),\end{aligned}$$

jossa a_i, a_j, $f_i(x,y)$ ja $f_j(x,y)$ ovat mielivaltaisesti valittuja vakioita ja (samankokoisia) kuvia. Kaavan ensimmäinen rivi määrittelee ominaisuuden nimeltä *additiivisyys* ja toinen *homogeenisyys*. Lineaarisuus määritellään tietysti vastaavasti summana, kun termejä on enemmän kuin kaksi.