

Uusi biosignaalien tallennusformaattistandardi

Alpo Värri

Signaalinkäsittelyn laitos, Tampereen teknillinen korkeakoulu, PL 553, 33101 Tampere

Johdanto

Monilla lääketieteellisillä mittalaitteilla kerätään potilaista mittaustietoa, joka tallennetaan digitaalisessa muodossa. Digitaalisten kuvien tallennukseen, varsinkin radiologiassa, on vakiintunut DICOM-standardi, mutta biosignaaleille ei ole toistaiseksi muodostunut yhtä yleistä tallennusmuotoa. Sellainen olisi kuitenkin tarpeen, kun terveystietomuksen liitteitäkin halutaan tallentaa digitaalisessa muodossa ja välittää eri valmistajien järjestelmien välillä.

Euroopan standardoimisjärjestön CEN:n teknillisen komitean TC251 (Health Informatics) piirissä käynnistyi vuonna 1994 työ, jossa laadittiin oliomalli mm. tehohoidon ja neurologian mittauksille ja tiedonsiirrolle (ENV 13734). Tässä yhteydessä koottiin myös mittava nimikkeistö mahdollisista biosignaaleista ja -mittauksista, mittauselektrodipaikoista, mittalaitteista ja mittayksiköistä ja näille annettiin yksikäsitteiset koodit. Tämä työ loi pohjan myös yhteentoimivan biosignaalien tallennusformaatin laatimiselle. Formaattia kehitettiin työnimellä File Exchange Format for Vital Signs (FEF) [1].

Vaatimuksia tallennusformaatile

Tallennusformaattistandardin laatiminen sai myös alkunsa vuonna 1994 kartoittamalla silloiset olemassaolevat tallennusformaatit ja niiden ominaisuudet. Havaittiin, ettei mikään olemassa ollut formaatti täytä kaikkia niitä vaatimuksia, joita laajasti käyttöön otettavalta formaatilta edellytetään. Vaatimuksia ovat mm. tuki binääriselle tiedontallennukselle (ASCII-pohjaiset formaatit rajautuvat näin pois), laaditun biosignaalitiedoston eri kohtien nopea haku riippumatta tiedoston koosta (joka voi olla parisataakin megatavua), eri näytteenottotaajuuksien tuki, mahdollisuus yhteenliittyvien mutta erillisten mittausseSSIoiden tallennukseen samaan tiedostoon, mittauksessa käytettyjen laitteiden identifioitavuus, mittaustietojen annotoijien identifioitavuus ym. Tallennusformaatin tulisi myös olla mahdollisimman yksinkertainen toteuttaa, jotta se voitaisiin ottaa nopeasti käyttöön pienissäkin yrityksissä, jotka kehittävät alalle laitteita ja ohjelmistoja.

Tallennusformaatin toteutus

Formaatista laadittiin oliomalli, joka pohjautuu esistandardiin ENV 13734 laajentaen sitä tarpeellisissa kohdissa. Formaatti on hierarkkinen ja myöhemmin laajennettavissa, sillä jokainen osa noudattaa rakennetta, joka alkaa osan yksilöllisellä tunnisteella, sisältää

seuraavaksi osan pituuden ja vasta sen jälkeen itse tallennettavan tiedon. Uutta formaatin määrittelyssä on mm. se, että aiemmin missään biosignaaliformaattissa ei ole erotettu tiedon abstraktia esitystapaa sen konkreettisesta esitystavasta. Tässä formaatissa on abstraktin esitystavan lisäksi määritetty yksi enkoodaustapa (ASN.1/BER) mutta toinen enkoodaustapa voidaan ottaa myöhemmin käyttöön, mikäli nähdään tarpeelliseksi. Standardoidun ja aikaisempia laajemman nimikkeistön käyttö on myös edistysakseli yhteentoimivuuden saralla.

Tiedosto, joka on tässä formaatissa (FEF-tiedosto), koostuu kuvan 1 mukaisista osista. Ensimmäinen tunniste (Preamble) sisältää FEF-tiedostoille oman tavusekvenssin, formaatin versionumeron, enkoodaustavan koodin ja tiedon, käytetäänkö yli 8-bittisten kokonaislukujen tavujärjestyksenä vähiten vai eniten merkitsevä tavu ensin järjestystä. Valmistajakohtainen osa (Manufacturer Specific Section) voi sisältää laitevalmistajan omaan käyttöön tallentamia tietoja, jolle formaatissa ei määritetä standardin mukaista esitystapaa. Tätä osaa ei välttämättä tarvita. Tässä osassa on myös valmistajan ainutlaatuinen oliotunniste (ISO 8824 object identifier tai ISO 11578 universal unique identifier), jota voidaan hakea ISO:n alaiselta organisaatiolta. Organisaation tiedot, kuten rekisteröinnin tehnyt sairaala ja mukana olleet henkilöt voidaan tallentaa omaan lohkoonsa (Health Care Provider Section). Potilaan yleistiedot, kuten nimi, syntymäaika, pituus jne. voidaan tallentaa omaan lohkoonsa (Patient Demographics Section). Tätä seuraavat Session Test –nimiset osat, joihin varsinainen mittaustieto on säilötty. Normaalisti niitä on vain yksi, mutta monimutkaisissa tiedonkeruutilanteissa voidaan joutua käyttämään useampiakin. Näitä seuraa Session Notes –niminen osa, johon voi tallentaa vapaamuotoisia kommentteja tehdyistä rekisteröinneistä.

Tunniste	Valmistaja-kohtainen osa	Organisaation tiedot	Potilaan yleistiedot	Session Test *	Session Notes *
----------	--------------------------	----------------------	----------------------	----------------	-----------------

Kuva 1. FEF-tyyppisen tiedoston pää rakenne. Poikkiviiva lohkon oikeassa yläkulmassa tarkoittaa, että kyseistä lohkoa ei ole tiedostossa pakko olla, jos sille ei ole tarvetta. * lohkoissa tarkoittaa, että kyseisiä lohkoja voi olla useampi kuin yksi.

Session Test –lohko voidaan jakaa kuvassa 2 näkyviin osiin. Mittauslaitteiston kuvausosassa (Medical Device System Section) esitellään kaikki rekisteröintiin liittyneet laitteet, joilta on tullut digitaalista mittausdataa kyseiseen tiedostoon. Mikäli mittauksen aikana mittauslaitteiston kokoonpano muuttuu, on aloitettava uusi Session Test -lohko. Näin voi tapahtua esimerkiksi tehohoitoympäristössä, jossa potilaaseen kiinnitetään uusi mittalaite. Vapaaehtoisessa multimediaosassa (Multimedia Section) voi olla digitaalisia kuvia, videopätkiä tai ääntä. Nämä voivat sijaita joko itse tiedostossa tai niihin voi olla vain viittaus, esimerkiksi hyperlinkki. Varsinaiset mittaustiedot sijaitsevat osassa nimeltä Session Phase. Niitä voi olla useampia kuin yksi, mutta esimerkiksi normaaliin koko yön unirekisteröintiin riittää yksinkertaisuuden vuoksi vain yksi Session Phase.

Mittauslaitteiston kuvaus	Multimediaosa	Session Phase *
---------------------------	---------------	-----------------

Kuva 2. Session Test –lohkon sisäinen rakenne.

Session Phase –lohkon sisäinen rakenne näkyy kuvassa 3. Ensimmäisenä sijaitsee tietorakenteen kuvausosa. Se jakautuu jatkuvien biosignaalitallenteiden (esim. koko yön mittainen katkeamaton EEG-signaali), lyhytaikaisten biosignaalitallennusten (esim. herätepotentiaalisygnalaali 1 sekunnin ajalta), jakaumamittausten (esim. EEG:n spektritieto), numeeristen mittausten (esim. lämpötilamittaus satunnaisena ajankohtana, nimenomaan ei tasavälein) ja erilaisten tapahtumakoodien (esim. QRS-kompleksin ajallinen paikka) tietorakenteiden kuvausosiin. Kuvausosa sisältää viittauksen tiedon tuottavaan laitteeseen, biosignaalilajin, mittayksiköt, anturin sijaintipaikan, tietotyypin yms. tiedot. Jatkuvien biosignaalitallenteiden osalta se kertoo myös, kuinka monta näytettä kultakin signaalikanavalta on yhdessä tietolohkossa ja missä järjestyksessä eri biosignaalit siinä ovat. Tietorakenteen kuvausosaa seuraavissa mittaustietojen osissa sijaitsevat itse mittaustiedot ilman, että mittayksiköitä yms. täytyy toistaa jokaisen mittauksen yhteydessä. Mittaustietolohkoja voi olla useita, mikä mahdollistaa sen, että eri tyyppisiä tietoja, kuten jatkuvia biosignaaleja ja tapahtumakoodeja voi tarvittaessa tallentaa lähes sitä mukaa kuin niitä ilmaantuu. Tällä tavoitellaan tiedostoformaatin soveltuvuutta reaaliaikaiseen tiedonkeruuseen, jolloin valmistajan ei välttämättä tarvitsisi toteuttaa lainkaan omaa sisäistä ja muiden laitteiden kanssa epäyhteensopivaa tiedostoformaattiaan.

Tietorakenteen kuvausosa	Mittaustietojen osa *
--------------------------	-----------------------

Kuva 3. Session Phase –lohkon sisäinen rakenne.

Jatkuva-aikaiset biosignaalit tallennetaan samanlaisissa lohkoissa kuin EEG-maailmasta tutussa EDF-tiedostoformaattissa [2]. Tällöin on periaatteessa mahdollista, että olemassaoleva EDF-tiedosto muutetaan FEF-formaattiin vaihtamalla biosignaalilohkojen edessä oleva kuvaileva alkuosa, nk. header-osa. Käytännössä tilanne ei kuitenkaan ole näin yksinkertainen, sillä muuttajan täytyy olla tarkkana muunnoksessa, jotta signaalien tyypit, anturien paikat, mittayksiköt jne. noudattavat tarkasti FEF-määrittelyn käytössä olevaa nimikkeistöä.

Esistandardi ENV 13734 sisältää FEF-formaatin käyttämän nimikkeistön ja koodiston mutta se on liitetty mukaan myös FEF-dokumenttiin liitteeksi A, jotta standardi voitaisiin toteuttaa tarvitsematta hankkia muita dokumentteja sen lisäksi. Koska koodisto on tarkoitettu reaaliaikakäyttöön, sen toteutuksessa pyrittiin tehokkuuteen rajaamalla koodin pituus kuuteentoista bittiin. Tämä katsottiin riittäväksi, koska koodialuetta voidaan uudelleenkäyttää kontekstiriippuvasti. Toisin sanoen koodi 2376 voisi merkitä sekä EMG-signaalia että mittayksikköä hehtaari kuin myös EEG-elektrodipaikkaa P4, koska

konteksti määrää, että tietyistä kohdista haetaan nimenomaan signaalin tyyppikoodia eikä mittayksikön koodia.

Laitevalmistajilta, jotka väittävät yhteensopivuutta kyseiseen standardiin, edellytetään dokumentaatio (Conformance statement), joka sisältää selvityksen siitä, miten mm. vapaaehtoiset osat on toteutettu. Näin toinen laitevalmistaja tai muu FEF-formaatissa tallennettujen tietojen käyttäjä voi ennakoita, mitä tietoja kyseinen laitteisto pystyy käsittelemään tai tuottamaan.

Pohdinta

Tallennusformaatin kehitystyön aikana tuli ottaa huomioon monia erilaisia näkemyksiä konsensuksen muodostamiseksi. Jos jokin erityisen huonoksi havaittu ehdotus haluttiin torjua, siihen tarvittiin hyvät perustelut. Osoittautui vaikeammaksi karsia formaatista ominaisuuksia kuin ottaa kaikki vaatimukset huomioon. Siksi yksinkertaisuudesta jouduttiin tinkimään kattavuuden kustannuksella. Standardidokumentin pituus onkin melkein 140 sivua ilman nimikkeistöä.

Aika näyttää, miten alan teollisuus ja tutkijapiirit tulevat ottamaan uuden formaatin vastaan. Sen leviämistä helpottaisivat vapaaseen jakeluun päästettävät apuohjelmat, joiden avulla formaattia voitaisiin lukea ja kirjoittaa ja sillä tallennettuja biosignaalitiedostoja voitaisiin katsella. ASN.1 on vielä vieras monelle bisignaalien parissa tutkimustyötä tehneelle tutkijalle ja voi osoittautua kompastuskiveksi osalle ohjelmoijista. Sen käyttöön on kuitenkin saatavissa ohjelmakirjastoja, joita käyttämällä formaatin toteutusta voidaan nopeuttaa.

Viitteet

[1] File Exchange Format for Vital Signs, First Working Document,
<http://www.cs.tut.fi/~varri/tc251/pt40/>

[2] B. Kemp, A. Värri, A.C. Rosa, K.D. Nielsen, J. Gade, "A simple format for exchange of digitized polygraphic recordings," *Electroencephalography and Clinical Neurophysiology*, Vol. 82, pp. 391-393, 1992.