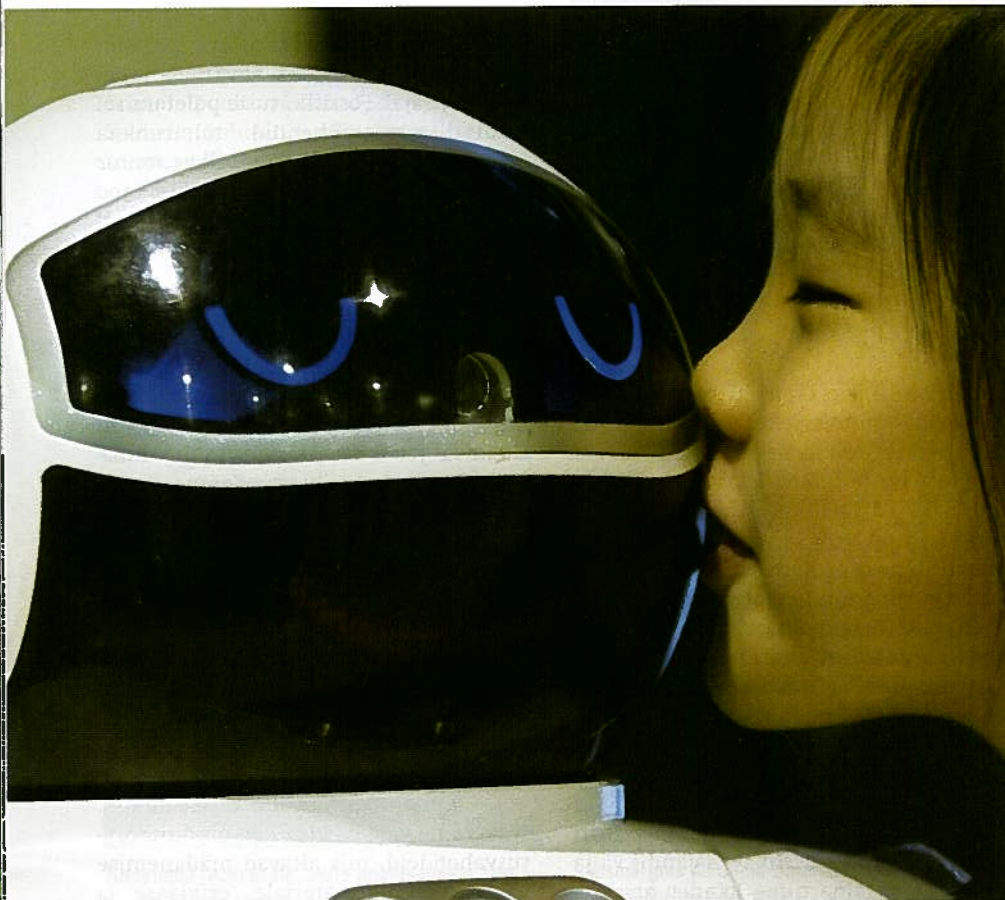


Inimene on alati püüdnud aru saada, kuidas maailm toimib. Kaua aega on uuritud siiski peamiselt maailma enese ümber ning reegleid, mis selle maailma liikuma panevad, et neid siis enese huvides ära kasutada. Ent ka inimese enda ja tema vaimu uurimine ei ole enam tabuteema ning üha enam püütakse mõista, kuidas "töötab" inimene ise. Püüame meiega heita pilku sellele teadmisele ja uurida, millist osa sellest on püütud järele teha arvutis.

MASINAD NÄGIJAKS

TOOMAS KIRT, MADIS LISTAK



Lõuna-Korea intelligentne robot Irobi 24. veebruaril 2004 reklaamüritusel ühes Souli kaubamajas. Irobi oskab näiteks õpetada inglise keelt ja suudab hoida videokaamera abil silma peal majas toimival. Irobi peaks müügile tulema juba tänava kevadel, hinnaks 3400 dollarit. Lisaks kaugjuhtimispuldile peaks Irobi mõistma ka inimhäälel öeldud käsklusi. Foto: Kim Kyung-Hoon/REUTERS

Inimene ammutab 70 protsenti infot ümbritseva maailma kohta silmade kaudu. Et robot saaks inimesele kasulik olla, peab ta samuti õppima nägema, ehk teisisõnu – olema võimeline tõlgendama visuaalset maailma inimesega sarnasel kombel.

Masinnägemine on üks tehisintellekti valdkondi, mille eesmärgiks on õpetada masinad masina või siis inimese kombel nägema. Aga miks just inimese moodi? Kas masin ei võiks näha ainult masina moodi?

Inimene elab suuresti enda loodud tehiskeskkonnas. Selleks et robotid oleksid niisuguses keskkonnas edukad ja inimesele kasulikud, peavad nad olema võimelised seda visuaalset maailma inimesega sarnasel kombel tõlgendama. Vähe sellest – erinevad kultuurid käsitlevad visuaalseid sümboleid ja värve erinevalt ning seega peavad ka robotid inimkultuuri erinevustest aru saama. Kujutuvastus sünteesibki teadmisi mitmetelt aladelt, nagu näiteks psühholoogia, bioloogia, filosoofia, statistika, signaalitöötlamine, geomeetria, modelleerimine, informaatika.

Järgnevalt mõningaist nägemise realiseerimise võimalustest arvutikeskkonnas. Kuna aga loodus on selle probleemi juba lahendanud ning inimesel ja loomadel on nägemisvõime välja arenenud, peatume esmalt mõningail inimnägemise külgedel, et oleks võimalik tuua paralleele selle keerulise protsessi kohta looduses ning arvutis. Keeruliseks muudab probleemi asjaolu, et nägemine ei ole ainult keskkonna fotografeerimine ja kahe pildi võrdlemine. Nägemine on ka pildil olevate objektide äratundmine ja nende tähenduse andmine. Nägemine on õpitud protsess.

INIMAJU TUGINEB TÖENÄOSUSELE

Evolutsiooni käigus on loomadel välja arenenud mitmesugused meeleorganid, millega aistitakse nii sisemisest kui ka ümbritsevast keskkonnast pärit ärritusi.

Visuaalse maailma tutvustamist arvutile nimetatakse masinnägemiseks, arvutinägemiseks või ka kujutuvastuseks, inglise keeles *Machine Vision*, *Computer Vision*, *Pattern Recognition*. Kõik need valdkonnad tegelevad sama probleemi eri tahkudega.

Erinevatele ärritustele reageerimine tagab liigile piisava informatsiooni keskkonnas toimetulekuks. Inimene tunneb näiteks valu ja maitseid, ta kuuleb ja näeb, tal on arenenud puutetundlikkus. Sensorsete protsessidega aistitakse esemete ja nähtuste üksikomadusi. Sellele järgneva protsessi tulemiks on taju – tegelikkuse esemete ja nähtuste vaimne esitamine. Esimeses etapis püütakse seega saada võimalikult hea kirjeldav pilt maailmast, seejärel toimub selle pildi töötlemine ja üksikkomponentide liitmine, et ära tunda olulisi objekte. Nägemine on protsess, mille käigus luuakse välise keskkonna piltidest kirjeldus, mis on vajalik vaatajale ja mis ei ole segi paisatud ebaolulise info tõttu.

Inimese poolt vaadeldav maailm on kolmemõõtmeline ning ka mõttes on võimalik asju ette kujutada kolmemõõtmelisena. Ent hoolimata sellest, et inimene suudab mõttes luua objektist kolmemõõtmelise kujutise, ei ole tal võimalik näha objekti kolmemõõtmeliselt, sest inimene ei näe ju objekti taha. Selle probleemi lahendab aju oma kujutlusvõimega ning inimene, kasutades varasemat objekti kirjeldust, loob mõttes kujutluse kolmemõõtmelisest objektist, ennustades, mis võiks olla objekti tagaküljel, kuhu ta ei näe.

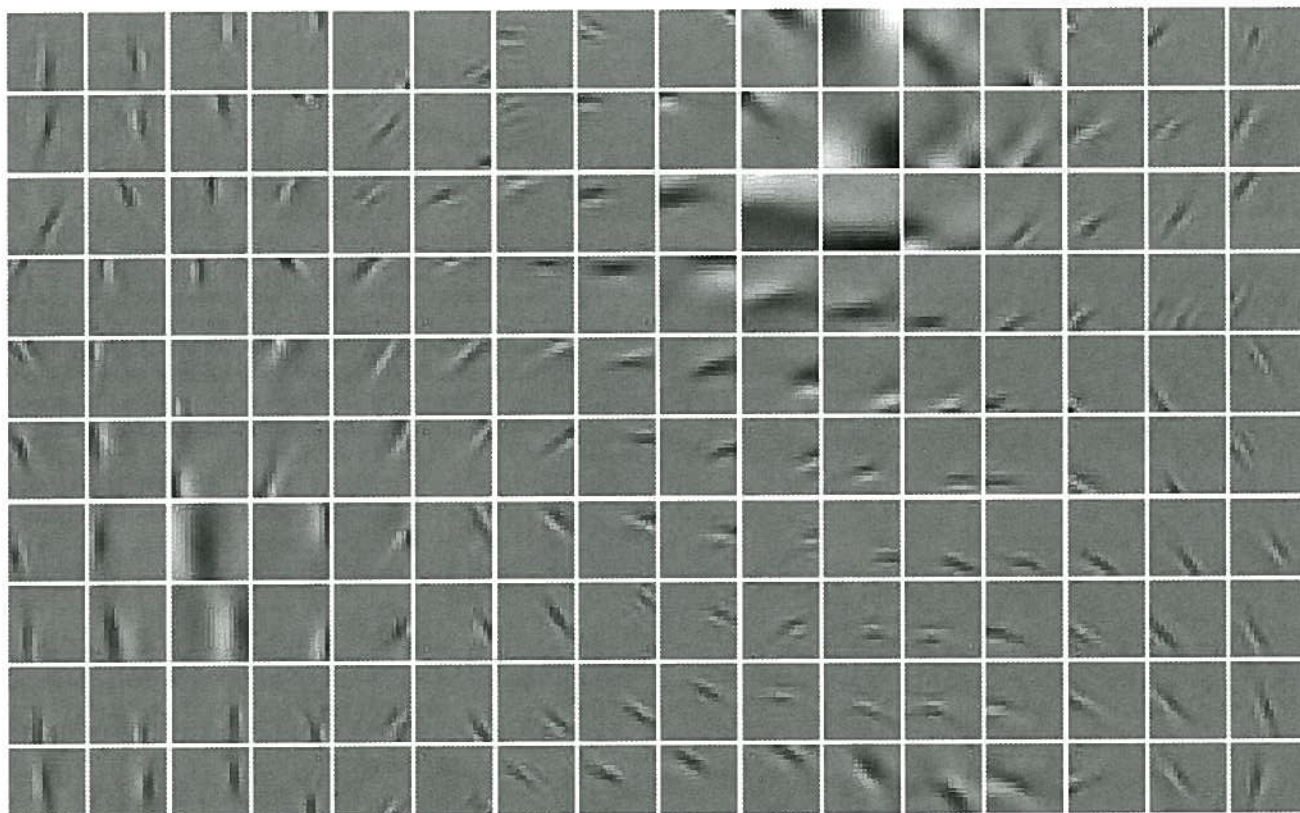
Kuigi inimene peab toime tulema kolmemõõtmelises ruumis ja suudab seda endale ette kujutada, on nägemise tinglik ulatus 2,5. Silma võrkkestal projekteerub kahemõõtmeline pilt ning vaadeldavale objektile lisatakse kaugusmõõde, mille võimaldab saada kahe silma abil loodav stereopilt – silmade omavaheline nurk annab ajule teada, kas objekt paikneb lähedal või kaugel.

Kuidas siis ikkagi tekib tunne, et see asi, mis on mu ees, on arvuti, millel ma parasjagu seda artiklit kirjutan? Põhimõtteliselt võib eeldada, et erinevate objektide äratundmine toimub hierarhiliselt – algselt kirjeldatakse objekti elementaarseid üksikosi, üksikosade kooslused annavad ettekujutuse tervikust ja see seotakse mõistega. Nii nagu sõnad ja laused pannakse kokku üksikutest tähtedest, nii pannakse ka tajutav objekt kokku algsetest üksikelementidest. Kõige raskem ülesanne ongi luua taoline tähestik, millega oleks võimalik kogu tajutavat maailma võimalikult hästi kirjeldada. Samuti puudub meil täna arusaamine, kuidas peaks välja nägema protsess, mis suudab tähestiku abil äratundmist vajavad objektid kokku panna.

Inimese ajus on esmased nägemisalad, mille rakud ärrituvad vaid teatava suuna, asukoha ja kujuga joonte olemas-

olul teatavas nägemispiirkonnas. Teatavate asukoha, kalde ja kujuga joonte kombinatsiooniga on võimalik nähtavat objekti kirjeldada ning seejärel on inimene võimeline seda objekti ära tunda. Lihtne näide: kui kaks horisontaalset ja kaks vertikaalset ühepikkust joont on otsapidi koos, siis on võimalik seda objekti ära tunda ja otsustada, et see on ruut. Seniste uuringute järgi võib eeldada, et kirjeldatud nägemisalad tekiavad inimese ajus pärast sündi, kui laps hakkab tutvuma ümbritseva keskkonnaga. Soome teadlased Aapo Hyvarinen ja Erkki Oja on modelleerinud samasuguse, inimajus toimuva esmaste nägemisalade tekitamise protsessi. Modelleerimise käigus eraldasid nad kümnest tuhandest juhuslikust väikesest pildi-fragmendist jooned, mis on vajalikud pildi iseloomulike omaduste kirjeldamiseks. Soomlaste poolt piltidest eraldatud jooned, mis sarnanevad inimaju nägemisaladega, on toodud joonisel.

Nägemine ei ole üksüheste vastavuste otsimine ajus, pigem püüab aju arvata oletada, mis objekt parasjagu vaateväljas on ehk teisisõnu – inimene kasutab objektide äratundmisel palju tõenäosust. Selle asemel, et objekti korralikult ja pikaajaliselt mõtta ning selle omadusi tundma õppida, loob aju hüpoteese, mis



Piltidest eraldatud sõltumatud komponendid. Soome teadlased Aapo Hyvarinen ja Erkki Oja on üritanud modelleerida inimajus toimuvat esmaste nägemisalade tekitamise protsessi – joonisel ongi toodud soomlaste poolt piltidest eraldatud jooned, mis sarnanevad inimaju nägemisaladega.

objektiga tegu on. Kui hüpotees peab paika, siis tajubki inimene objekti sellena, mis see on. Kui hüpotees ei pea paika ja objekt ilmnepd uued tunnused, mis eelmise hüpoteesi kummutavad, siis püüab aju uute, lisandunud tunnuste alusel objekti ära tunda.

Seega näeb inimene kõige tõenäolisemat välise maailma olukorda. Üldjuhul on aju võimeline seda ära arvama, kuid on ka mitmeid visuaalseid efekte, mida aju tõlgendab pisut ekslikult.

MASIN OSKAB JÄÄDVUSTADA

Ümbritsevat keskkonda jäädvustada me juba masinate abil suudame, selleks piisab lihtsast fotoaparaadist või videokaamerast. Elementaarne pildituvastuse süsteem oleks seega kahe pildi võrdlemine. Kui need täpselt kokku langevad, on selge, et tegemist on sama objektiga. Probleemi muudab keerulisemaks asjaolu, et kahjuks ei ole n-ö samasse klassi kuuluvad objektid alati ühesuguse kujuga. Ja isegi kui on tegemist sama objektiga, ei paikne see alati täpselt samas positsioonis – on teismoodi valgustatud

ja erinevad ka paljud muud tegurid, mis mõjutavad pildi kvaliteeti.

Masinnägemise saab jaotada kaheks osaks: madalam tasand (*low-level vision*) ja kõrgem tasand (*high-level vision*). Madala taseme nägemisprotseduurid tegelevad värvide ja visuaalse müra töötlemisega, pildi teravustamise ja piirjoonte leidmisega, valguse, varju ja värvitoonide ühtlustamisega. See valdkond on otseselt välja kasvanud signaalitööstusest ja hästi uuritud. Sama tehnika on kasutusel ka näiteks video- ja fototöötusel. Meetodeid eristatakse ka selle järgi, kas töödeldakse ainult üksikuid pilte või pildijadasid (videoid).

Kõrgema taseme meetoditel on aga rohkem kokkupuutepunkte modelleerimise, visualiseerimise ja ka psühholoogiaga. Paljud meetodid modelleerivad kõigepealt kolmemõõtmelise geomeetria mudeli abil objekti, loovad siis projektsioone ja võrdlevad neid projektsioone nähtava pildiga, enne kui otsustavad, mida nad näevad. Nii toimivad näiteks mõned inimnäo või satelliidipiltide tuvastusalgoritmid.

ARVUTI POLE INIMAJU

Selleks et pildi inimlikku sisu arvutile kuidagi kirjeldada, on vaja pilt algosadeks jagada. Seda võib teha piirjoonte, sama värvi piirkondade või mõne muu tunnuse, näiteks tekstuuri alusel. Niisugune tegevus on kõrgema taseme nägemisprotseduuride eelduseks ja samas ka üks osa neist.

Kõrgema taseme meetodid tegelevad seega pildist aru saamisega. See on tänapäeval veel lõpuni läbi töötamata valdkond ning leidub palju mitmesuguseid teooriaid ja meetodeid, kuidas pilti tõlgendada. Me ei saa siin ka täiesti inimest kopeerida, sest esiteks me ei tea lõplikult, kuidas inimene seda teeb, ja teiseks on meie arvutid teistsuguse ehitusega kui inimaju.

Kõrgema taseme meetodid püüavad pilti ja selle osi analüüsides leida miniguid unikaalseid omadusi, mis üheselt seoksid vaadeldava objekti mingi arvutile teada oleva objektiga. Selleks võib olla näiteks pildipunktide mingi statistiline seaduspära, värv, tekstuur või ka piirjoontest moodustuv geomeetiline



Üks valdkond, kus masinnägemine inimesele juba praegu appi saab tulla, on veekogu seisundi analüüsimine taimestiku alusel. Pildidel on näha mõned Eesti rannikumere põhja piltide töötlemise etapid: vasakpoolsetel on normaliseeritud kujutis, keskmistel on eemaldatud enamik vett ja merepõhja, parempoolsetel on ruudukestega märgitud tuvastatud taimestik.



kuju. Suureks probleemiks on tagada nende omaduste sõltumatus valgustingimustest, vaatenurgast ja objekti kaugusest. Muutub ju eseme värv sõltuvalt kellaajast ja valgusallikast, eseme mõõdud sõltuvad selle kaugusest, eseme tähendus aga teiseneb vahel sõltuvalt sellest, kas ese on püsti või pikali. Objekt võib olla sageli ka osaliselt varjatud.

Üldisemaks probleemiks on aga asjaolu, et inimese nägemine on õpitud tegevus ja ei allu seepärast hästi rangetele matemaatilistele meetoditele. Inimese aju vahel lihtsalt "parandab" pilti, mis silmadesse jõuab. Teiseks probleemiks on see, et looduslikel objektidel, erinevalt masinaosadest ja inimkäte poolt loodud ehitistest, puuduvad tihti ühesed geomeetriselised või statistilised tunnused, nende äratundmist saab korraldada vaid mitmeid tunnuseid üheaegselt arvesse võttes. Viimati nimetatud asjaolu ongi tihti veelahkmeks masina moodi nägemise ja inimese moodi nägemise vahel. Masina moodi nägemise meetodeid kasutatakse kontrollitud keskkonnas, kus masinal pole üllatusi oodata. Kontrollimata keskkond aga kujutab endast masina jaoks suurt probleemi. Sellest asjaolust tulenevalt ei annagi ükski puhtalt statistikat või geometriat kasutav meetod kontrollimatus keskkonnas head tulemust. Siin paikneb ka praegu uurimistöo raskuspunkt ja tuhandeteadlased kogu maailmas otsivad probleemile lahendusi.

Masinnägemine on viimased aastakümned kasutust leidnud enamasti vaid teaduslaborites, suurettevõtetes või sõjanduses. Äsja alanud kümnend toob siin kindlasti muutusi ja näeme rohkem rakendusi igapäevaelus. Tehisintellekt koos masinnägemisega ja robotid muutuvad meie elus sama tavaliseks kui mobiiltelefonid või personaalarvutid täna.

Eestis tegeletakse selles valdkonnas veetaimede tuvastamise probleemiga. See on vajalik merekeskkonna seisundi hindamiseks. Praegu kulub selleks palju aega ja inime tööjõudu.

Niisiis on vaja arendada arvutite nägemisvõimet, kaasamaks neid üha enam seni inimese pärusmaaks olnud tegevustesse. ■

TOOMAS KIRT (1971) on TTÜ doktorant. Uurimisteenaks andmeanalüüsi meetodid ja neurovõrgud.
MADIS LISTAK (1966) on TTÜ doktorant. Uurimisteenaks masinnägemine ja keskkonnaseire robotid.

LOE VEEL

Steven Pinker. *How the Mind Works*. Penguin Books, 1999.
Talis Bachmann. *Rait Maruste. Psühholoogia alused*. Ilo, 2003.

MASINNÄGEMISE KASUTUSVALDKONNAD TÄNAPÄEVA MAAILMAS

- **Satelliidipiltide töötlemine** – ilmaennustamine, keskkonnaseire (merereostuste jälgimine ja avastamine), metsade ja põldude seisukorra hindamine aerofotode alusel, luure, kartograafia.
- **Jälgimis- ja turvasüsteemid** – näevad masinad ehk arvutid, kes oskavad lugeda autode numbrimärke. Näiteks aitavad arvutid koguda Londonis autoomanikelt kesklinna sõitmise eest maksu, samamoodi tahetakse Saksamaal teemaksu korjama hakata.
- **Turvasüsteemid** – arvutid otsivad lennujaamades terroriste näopiltide järgi, pääslasüsteemides vaatab arvuti näpujälgi ja tuvastab selle kaudu isikuid. Ameerikas on mõnes poes kasutatud seda süsteemi kassas kauba eest maksmiseks.
- **Automatiseeritud dokumentitöötlus** – skannerid oskavad dokumentifoto arvutitekstiks muuta.
- **Tööstusrobotid ja automatiseeritud tootmisliinid** – arvutinägemist kasutatakse detailide sorteerimiseks ja kvaliteedi kontrolliks.
- **Mood** – katsetatakse süsteeme, mis oskavad visuaalselt võtta inimkehalt vajalikud mõõdud, et ömmelda ideaalselt sobivad riided.
- **Meelelahutustööstus** – filmides kasutatakse sünteetilisi ehk tehisiikke, ainult arvutis eksisteerivaid näitlejaid. Need luuakse päris inimeste ja loomade järgi, muutes fotod ja videod arvutimudeleiteks. Sama tehnikat kasutatakse ka arvutimängude loomisel. Paljud kasutatavad tehnikad on siin samad kui arvutinägemises.
- **Meditatsioon** – arvutid analüüsivad röntgenipilte ja abistavad diagnoosimisel. Loomisel on ka esimesed elektroonilised silmad pimedatele nägemise taastamiseks – ka sellisel juhul toimub esmane pilditöötlus arvutis ja alles seejärel suunatakse tulemus aju.
- **Põllumajandus** – arvuti aitab näiteks automatiseeritud kasvuhoones eristada küpsed tomatid toorestest, hoida traktori täpselt vao peal, et teha vaheltharimist, eristada umbrohtu kultuurtaimedest.
- **Teadus** – mikroskoobipildi töötlemine DNA-uurimises, aga ka asteroidide ja komeetide leidmine fotodelt.
- **Internet** – masinnägemine võimaldab teha päringuid internetist, näiteks on võimalik koostada päring, mis otsib ainult punase roosi pilte.
- **Arhiivindus** – võimalus indekseerida videoarhiive, et hiljem oleks võimalik automaatselt otsida huvipakkuvat infot.
- **Sõjandus** – relvade (suurtükid, tiib- ja õhutorjakeretid) juhtimissüsteemid. Aga ka robotrelvade navigatsioon.
- **Liiklus** – suurt tööd tehakse selle nimel, et turule tuua automaatnavigatsiooni seadmed, mis lubaksid loobuda autojuhist või siis aitaksid ohuolukordi paremini lahendada.
- **Turism** – on loodud pihuarvutit töötavaid süsteeme, mis oskavad võõrkeelseid tänavasilte mõnda teise keelde tõlkida.
- **Masinnägemist kasutatakse ka kantavate arvutite (*wearable computers*) juures** – puusal kaasas kantava arvutisüsteemi koostisosaks on prillidele kinnituv kaamera ja seade, mis kaamerapilti silma projekteerib. Lisaks on loodud tarkvara, mis võimaldab ümbritsevast maailmast välja lõigata kõik visuaalsed reklaamid ja asendada reklaamidest kirjuks maalitud majaseinad millegi meeldivamaga – näiteks lillepiltidega.



Turvakaamera "silma" jälgimas toimuvat Ateena kiirteel 25. veebruaril 2004. Ateena suveolümpiamängudeks on muu hulgas plaanis paigaldada 1250 stantsioneeret turvakaamerat, mis paigutatakse 12-meetrise sammaste otsa.

Foto: Thanassis Stavrakis/AP Photo