

Makinak norberaren ahotsaz hitz egin dezan

Ana Galarraga Aiestaran · Elhuyar Zientzia

Inma Hernaez Rioja Bilboko Goi Ingeniaritzako Eskolako irakasle katedraduna da, eta Aholab ikerketa-taldearen sortzailea. Taldea 1995ean jaio zen, EHUren barruan, eta euskarazko hizketa automatikoa lantzen dute. Nahiko lan isila izan da, orain arte. Hain zuzen, azken proiektuak, [AhoMyTTS](#), izan dituen oihartzuna eta harrera hain izan dira handiak, ezen sistema eten egin behar izan zuten, tarte batez.

Arrakasta horren atzean, urte askotako lana dago, eta abiapuntuan, berriz, bizipen pertsonal

bat: Hernaezen ahizpak ahotsa galdu zuen, operazio baten ondorioz. Eta hari eta hura bezalakoiei komunikatzen laguntzeak bultzatu du bere ibilbidean. "Ahizpa izan da gure garapenen lehen erabiltzailea. Behin, ahizparen lagun min bati omenaldi bat egin zioten, eta omenaldi hartan ahizpak hitzaldi txiki bat eman zuen, aplikazioari esker. Aukera hori emateak sekulako indarra eman dit ikerketan aurrera egiteko".

Hastapenetatik gaur egunera, garapena nabarmena izan da. Orain dela 20 urte, ahots-sintesiko tek-

Nahikoa da erabiltzaileak ehun esaldi grabatzea, batzbesteko ahotsari egokitzeko berea. ARG.: Branislav Nenin/Shutterstock.com.





niketan kalitate ona lortzeko, pertsona baten grabaketa-ordu asko behar ziren. Adibidez, Sirik duen ahotsa lortzeko, 30 ordu inguru beharko zituztela esan du Hernaezek. Gainera, estudio batean grabatu behar zen, oso kalitate onarekin.

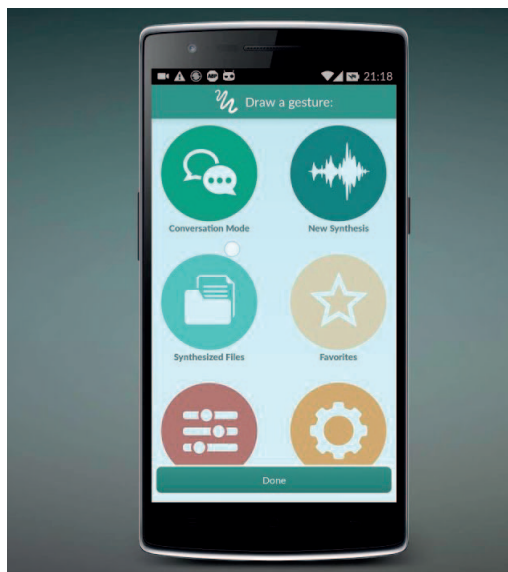
Teknika “moztu eta itsatsi” prozeduran oinarritzen zen. Ez ziren hitzak edo esaldiak mozten, baizik eta unitate txikiak: difonemak, trifonemak, bokalak edo kontsonanteak testuinguruarekin... Gero, testuinguruaren arabera, algoritmoak erabakitzen zuen zer unitate aukeratu datu-basetik. “Korpusean oinarritutako sintesia deitzen zitzaion. Oso garestia zen”, zehaztu du.

Erabilgarritasuna eta kalitatea helburu

Hurrengo urratsa 2002. urtearen inguruan etorri zen, sintesi estatistiko-parametrikoiari esker. Teknika horretan, lehenik, ahots-seinaleak parametrizatu egiten dira, eta, gero, parametro horiek elkar lotzen dira, tratamendu estatistiko baten bidez. Azaldu duenez, kalitatea ez da aurreko sistemarena bezain ona, baina askoz ere grabaketa gutxiagorekin baleko emaitza lortzen da. Gainera, estatistikoa izateak malgutasun handia ematen dio, eredu estatistikoa edozein lagini egokitu baitaioke.

Honela azaldu du prozesua. “Badugu batazbesteko ahotsa. Ahots hori ez da inorena, ahots eta hizlari askorekin egina dago. Ez dira edozein hizlari, hizlari profesionalak izaten dira; haien grabaketatik abiatuta lortzen da batazbesteko ahotsa. Gero, ahots hori edozein erabiltzaileraren ahotsera egokitu daiteke, eta, horretarako, nahikoa da ehun esaldiko lagin bat izatea”.

Ehun esaldi horietan, soinu-konbinaketa guztiak daude, eta grabaketak ordu erdi inguru hartzen du. Aurrerapausoa agerikoa da, beraz, asko errazten baitu edonork izatea bere ahotsera moldatutako ahots sintetikoa. “Berez, esaldi gehiago grabatuz gero, kalitate hobea lortzen da, baina ezin diogu hori baino ahalegin handiagoa eskatu jendeari”, dio Hernaezek.



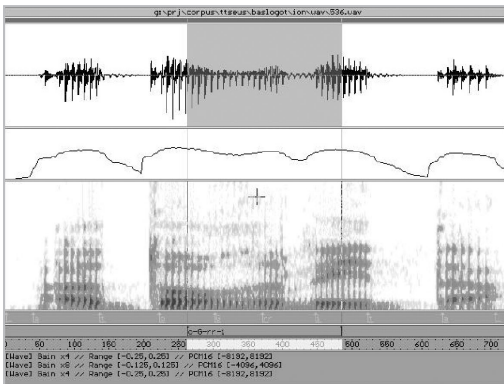
Norberaren ahots sintetikoa erabiltzeko aplikazioa erraz implementatzen da Android mugikorretan. ARG.: Aholab.

Beste muga batzuk ere baditu. Esaterako, ez da emozioak adierazteko gai: “Emoziodun korpus bat grabatuz gero, lor daiteke, baina, proiektu honetan, esaldiak neutroak dira, eta ez dugu egiten testuaren azterketa, emozioak bilatzeko. Hala ere, testuan jar daiteke, baina ez dugu egiten”.

“Lanean jarraitzen dute kalitatea hobetzeko (algoritmo berri bat dute) eta laborategitik erabiltzailearengana egiten den bidea laburtzeko”

Baina alde onak nabarmenak dira: sintetizagailuak ez du memoria hartzen, erraz implementatu dute Androiden, mugikorrean denbora errealean funtzionatzen du... Horrek guztiak aukera ematen dio edonori erabiltzeko egunerokoan. Eta hori zen hasierako helburua.

Izan ere, Hernaezek aitortu du gaur egun badaudela kalitate hobea ematen duten sistemak, bereziki, sare neuronaletan oinarritutakoak: “Egungo sintesiteknologia izugarri ona da. Nik erakutsiko nizkizuke esaldi batzuk, eta zu ez zinateke gai izango bereizteko sintetikoa den ala naturala. Baina ezingo genuke Androiden sartu, edo denbora errealean erabili”.



Garatu duten proiektua sintesi estatistiko-parametrikorearen teknikan oinarritzen da. ARG.: Aholab.

Laborategitik erabiltzailearengana

Eta koska hori da: egunerokoan, komunikatzeko ahots sintetikoa behar dutenek dauden sistemak erabili behar dituzte, eta horiek oso mugatuak dira. Hala, kasu askotan, emakumeek gizon-ahotsa erabili behar dute; edo hizkuntza gutxitue-

tako hiztunek, hizkuntza hegemonikoa; edo haurrek, Sirirena. Erabiltzailearen ahotsaren eta sintetizagailuak eskaintzen duenaren tartea txikitzea izan da Aholaben xedea.

Garatu duten sistema, beraz, erabiltzailearentzat ahalik eta erosoan eta errazen izan dadin saiatu dira. Hau da funtzionamendua: “Erabiltzaileak lehenik ehun esaldi horiek grabatzen ditu, euskaraz edo gaztelaniaz, eta automatikoki sortzen da bere ahots sintetikoa duen aplikazio bat. Segidan, mezu elektronikoa bidez jasoko du, eta nahikoa du hor klik egitea, bere Android mugikorrera jaisteko. Ahotsa sistemako ahots gisa grabatzen da; beraz, komunikatzeko gure aplikazioan ez ezik, beste aplikazioetan ere erabil dezake, adibidez, liburuak irakurtzeko e-reader askotan, edo Adoben. Eta denbora errealean komunikatu daitezke”.

Mugikortasuna oso mugatua dutenek, berriz, tableta eta palanka edo iris-irakurgailua erabiltzen dutela zehaztu du. Kasu horretan, Windowsen integratzen dute.

Hain zuzen, erabiltzaileen artean daude alboko esklerosi amiotrofikoa duten pertsonak. Berez, ahotsa galdu duen edonori laguntzeko helburua du proiektuak, baina galera bat-batekoa denean (iktus baten ondorioz, esate baterako), zaila da bere ahots sintetikoa lortzea. Alboko esklerosi amiotrofikoa dutenek, berriz, pixkanaka galtzen dute mugitzeko ahalmena. Hortaz, diagnostikoa jasotzen dutenetik, denbora dute grabaketa egiteko eta beren ahots sintetikoa sortzeko.

Azken hiruzpalau urtetan, berriz, laringektomia egin behar zitzaizen pazienteekin aritu dira, Biocrucesekin batera. "Laringektomizatutakoek ahots esofagikoarekin hitz egin dezakete, baina, elkarlan honen bidez, medikuek informazioa ematen zieten pazienteei, ebakuntzaren aurretik grabaketa egiteko. Hartara, ondoren bazuten aplikazioa, lehen asteetan ezin baitute hitz egin", azaldu du Hernaezek. Dioenez, elkarlan horrek bultzada handia eman dio proiektuari.

Orain, lanean jarraitzen dute kalitatea hobetzeko (algoritmo berri bat dute) eta laborategitik erabiltzailearengana egiten den bidea laburtzeko. Ume-ahotsa ere eskuragarri jartzeko ahaleginak egiten ari dira, baita beste hizkuntza batzuetan ere eskuragarri izan dezaten, adibidez, katalanez. Azken helburua: behar duten guztiek izan dezatela ahots sintetikoaren bidez komunikatzeko aukera, ahalik eta modu pertsonalizatu, natural eta samurrenean. ●



Erabiltzaileen artean, alboko esklerosi amiotrofikoa duten pertsonak daude. Mugikortasuna oso mugatua dutenek tableta eta palanka edo iris-irakurgailua erabiltzen dute. ARG.: Sasirin Pamai/Shutterstock.com.