

PERTEMUAN 15

SISTEM BERSENSOR GANDA

A. TUJUAN PEMBELAJARAN

Pada pertemuan ini akan dijelaskan mengenai sensor input, multimedia dan sistem multimedia, speech pada interface, non speech sound, pengenalan tulisan tangan, pengenalan gerak dan computer vision, aplikasi multimedia sistem. Setelah menyelesaikan pembelajaran pada pertemuan 15, mahasiswa mampu:

1. Menjelaskan sensor input, multimodal dan multimedia
2. Menjelaskan pemakaian speech pada interface
3. Menjelaskan pengenalan tulisan tangan, pengenalan gerak dan komputer vision
4. Menjelaskan pengertian text, hypertext dan hypermedia
5. Mengenali aplikasi multimedia system

B. URAIAN MATERI

1. Menjelaskan sensor input, multimodal dan multimedia

System multi sensor (beberapa sensor) memakai lebih banyak sensor / sensor secara bersamaan.

Contoh :

- a. Bunyi
- b. Bacaan
- c. Hypertext
- d. Kartun
- e. Film
- f. Gesture
- g. vision (mata)
- h. Dan lainnya

Digunakan dalam beberapa aplikasi: pada dasarnya cocok untuk yang menggunakannya dengan kebutuhan special, serta dalam realitas buatan (Virtual Reality).

Lima indra bisa dapat dipakai yaitu:

- a. Pandangan (mata)
- b. Suara
- c. Sentuhan
- d. Pengecapan
- e. Dan penciuman

a. Sensor Input

Di komputer, saluran visual digunakan sebagai saluran komunikasi utama, tetapi jika kita ingin menggunakan indra lain, kita harus mempertimbangkan relevansinya dan sifat informasi yang dapat dikirimkannya. Penggunaan suara merupakan area yang jelas untuk eksplorasi di masa depan. ucapan digunakan sampai batas tertentu di banyak antarmuka: sinyal akustik digunakan sebagai peringatan dan pemberitahuan, dan ucapan serta musik yang direkam atau disintesis juga digunakan. Saat ini, ada beberapa metode implementasi perangkat yang mampu menghasilkan rasa dan bau, oleh karena itu kedua area tersebut tidak didukung.

b. Sistem multimodal

Secara umum, sistem multimodal didefinisikan sebagai sistem khusus yang mendukung komunikasi dengan pengguna melalui berbagai modalitas, seperti suara, tulisan tangan, gerak tubuh, emosi, dll. Dalam proses interaksi, baik antara manusia maupun antara pengguna manusia dan sistem komputer, gaya dan gaya berperan. Metode menentukan jenis data yang akan dipertukarkan, sedangkan mode menentukan konteks di mana data diinterpretasikan. Oleh karena itu, jika kita memiliki pandangan yang lebih sistemik, multimodal dapat dianggap sebagai kemampuan suatu sistem untuk berkomunikasi dengan pengguna menggunakan berbagai jenis saluran komunikasi dan untuk mengekstrak dan mengirimkan makna secara otomatis.

Sistem multimodal itu sendiri digunakan kian dari 1 indera atau cara interaksi.

Contohnya :

Indera pengelihatan serta pendengaran: Pengolah kata dapat mengucapkan frase dengan menampilkan kata-kata di layar.

c. Sistem Multimedia

Secara umum multimedia adalah media yang menggabungkan dua unsur atau lebih media terdiri dari :

- 1) Teks
- 2) Grafis
- 3) Gambar
- 4) Foto
- 5) Audio
- 6) Video
- 7) Dan animasi secara terintegrasi

Sistem yang dipakai yaitu:

- 1) gunakan media yang berbeda untuk menyampaikan informasi.
- 2) Contoh: system pendidikan terkomputerisasi dengan menggunakan vidio, kartun, teks, serta gambar diam.
 - a) Berbagai alat komunikasi menggunakan bentuk interaksi visual.
 - b) Anda bisa juga menggunakan vokal, untuk berbicara serta tanpa berbicara: dua metode lainnya, sekarang memakai metode berbeda.

2. Menjelaskan pemakaian speech pada interface

Perkataan (speech) banyak ditambahkan dalam tampilan interface speech ini dalam interface berguna untuk keadaan:

- a. Tangan user yang sedang sibuk
- b. Mata user harus memperhatikan sesuatu
- c. Kondisi yang tidak memungkinkan menggunakan *keyboard*, misalnya, meja tempat user bekerja banyak tumpukan kertas.

a. Pengenalan Perkataan (*Speech Recognition*)

Banyak upaya telah dilakukan untuk mengembangkan sistem pengenalan suara, tetapi meskipun sistem komersial sekarang tersedia secara luas dan murah, keberhasilannya masih terbatas pada sistem

pengguna tunggal yang memerlukan pelatihan yang signifikan. Salah satu hambatan ketidaksuksesan tersebut adalah :

- 1) Kompleksitas bahasa
- 2) Derau latar belakang dapat mengganggu masukan, mengaburkan, atau mengubah informasi, sementara headphone dapat memetakan suara yang berlebihan atau tidak berarti dalam aliran informasi, mengulang, menghentikan, atau menggunakan suara "berkepanjangan", seperti "Ummm" dan "Err" untuk mengisi jeda pada suatu pembicaraan.
- 3) Variasi Individu, Orang-orang memiliki suara yang unik dan sistem berkinerja tinggi dikonfigurasi agar peka terhadap sedikit perbedaan nada dan frekuensi suara pembicara - speaker baru menawarkan refleksi sistem yang berbeda dan oleh karena itu tidak dapat bertindak.
- 4) Dialog daerah yang sangat berbeda dapat menyebabkan masalah pada sistem

Pada titik paling sederhana, beberapa sistem mengharuskan Anda untuk menghubungi nomor ekstensi, tetapi karena cincin telah menjadi universal, manfaat entri ini terhadap nomor tersebut dipertanyakan! Sistem lain menggunakan audio secara lebih aktif, terutama sistem informasi untuk pemesanan tiket pesawat.

b. Perkataan Buatan (Speech Synthesis)

Speech synthesis melengkapi pengenalan ucapan. Ide untuk dapat berbicara dengan komputer secara alami menarik banyak pengguna, terutama mereka yang tidak menganggap dirinya akrab dengan komputer, karena ini mencerminkan ekspresi normal dan alat komunikasi sehari-hari mereka. Namun, ada banyak masalah dengan sintesis dan pengenalan ucapan. Masalah tersebut adalah:

- 1) User sangat peka terhadap perbedaan nada bahasa, dan oleh karena itu, kami tidak mentolerir kekurangan pidato yang kompleks.
- 2) Membutuhkan lebih banyak kebisingan di lingkungan kantor atau penggunaan headphone, yang keduanya bisa menjadi harga yang sangat tinggi untuk membayar manfaat sistem.
- 3) Ucapan yang majemuk

Beberapa lingkungan aplikasi speech synthesis yang telah berhasil:

- 1) Speech synthesis menawarkan media aplikasi kepada tunanetra yang dapat mereka akses.
- 2) visual pengguna difokuskan di tempat lain, seperti peringatan di kabin dan, baru-baru ini, di mobil.

3. Menjelaskan jenis-jenis non-speech sound

Suara non-verbal dapat menawarkan sejumlah keuntungan. Suara non-verbal bisa diserap lebih cepat. Pidato bergantung pada bahasanya: sistem berbasis ucapan memerlukan kemampuan untuk menggunakan terjemahan untuk kelompok bahasa lain. Suara non-verbal dapat menggunakan fenomena penyesuaian pendengaran: suara latar diabaikan kecuali jika diubah atau dimatikan.

Audio non-verbal juga dapat digunakan untuk memberikan representasi kedua dari tindakan dan objek pada antarmuka untuk mendukung pose visual dan ketegasan pengguna. Eksperimen menggunakan navigasi pendengaran, menunjukkan bahwa sinyal pendengaran cukup bagi pengguna untuk mendeteksi hingga delapan target di layar dengan kecepatan dan akurasi yang wajar. Ini menunjukkan bahwa tidak ada alasan yang baik untuk mengabaikan peran suara pada antarmuka, karena bisa sangat kabur atau tidak akurat. Tetapi jenis suara tak terucapkan apa yang harus kita gunakan pada antarmuka? Ada dua alternatif: menggunakan suara alam dunia dan menggunakan suara yang lebih abstrak.

a. Soundtrak

Soundtrack adalah contoh bagus pengolah kata dengan antarmuka audio, dirancang untuk pengguna tunanetra. Elemen visual yang disajikan memiliki analog pendengaran, terdiri dari warna, dengan ucapan yang kompleks. Sebuah grid yang terdiri dari dua baris dan empat kolom adalah layar utama dari soundtrack.

Tabel 15.1 **Layar utama soundtrack**

File Menu	Edit Menu	Sound Menu	Format Menu	
Alert	Dialog	Document 1	Document 2	

Setiap sel memancarkan nada yang berbeda saat kursor berada di dalam dan dengan nada ini pengguna dapat bergerak di sekitar sistem. Mengklik sel membuat nama Anda berbicara dan memberikan informasi spesifik yang dapat mengarahkan pengguna yang tersesat atau bingung.

Soundtrack memungkinkan entri teks yang mengucapkan kata atau huruf saat disisipkan, dengan pengguna mengontrol klasifikasi komentar.

Soundtrack menyediakan solusi audio berbasis kata untuk visualisasi pengolah kata, meskipun hasilnya mungkin tidak terlalu konsisten dengan antarmuka visual.

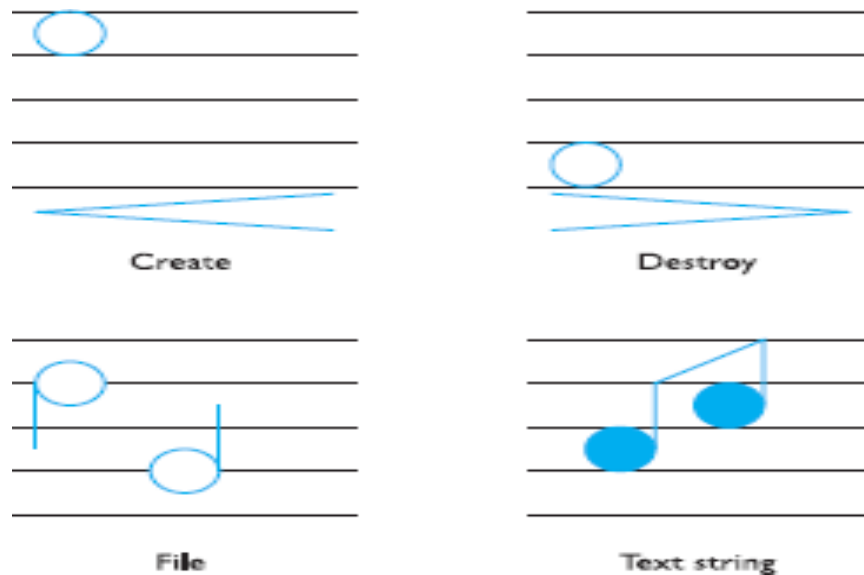
b. Auditory Icons

Ikon audio memakai bunyi natural untuk mewakili berbagai jenis objek dan tindakan di antarmuka. SonicFinder untuk Macintosh dikembangkan dari ide-ide ini untuk meningkatkan antarmuka dengan iterasi. Suara alam digunakan karena manusia mempelajari sumber dan perilaku suara, bukan suara dan nada. Misalnya, suara akan diidentifikasi sebagai pecahan kaca atau diklik sebagai tabung berlubang. Pengenalan ini sangat kompleks: Anda tidak hanya dapat menentukan sumber suara (misalnya, untuk menulis tabung), tetapi juga properti sumber suara (misalnya, apakah tabung berlubang atau padat).

Di SonicFinder, ikon audio digunakan untuk mewakili objek dan tindakan di desktop. Jadi, misalnya sebuah volume ditampilkan dengan suara yang mirip dengan kertas, dan ada sesuatu yang dibuang ke tempat sampah dengan suara pecah.

c. Earcons

Alternatif untuk menggunakan suara alam adalah dengan membuat suara buatan. Pilihan telinga menggunakan kelompok catatan yang terorganisir, yang disebut pola, untuk mewakili tindakan dan objek.



Gambar 15.1 Earcons

ini bervariasi dengan waktu, nada, warna, skala dan volume. Ada dua jenis headphone. Headphone komposit menggabungkan dekorasi yang berbeda untuk membuat tindakan tertentu, misalnya, untuk menggabungkan pola "kreasi" dan "gulungan". Seni telinga yang sudah dikenal menggambarkan suara telinga yang terdiri dari jenis yang serupa. Misalnya, kesalahan sistem operasi dan kesalahan sintaks adalah bagian dari keluarga "kesalahan". Overtones menawarkan pendekatan terstruktur untuk antarmuka desain suara, tetapi dapatkah pengguna mempelajari lebih lanjut tentang suara dan faktor apa yang mempengaruhi penggunaannya? Bukti menunjukkan bahwa orang dapat belajar mengenali alat bantu dengar dan bahwa elemen terpenting dalam membedakan antara warna suara yang berbeda adalah kualitas suara yang berbeda yang dihasilkan oleh instrumen dan suara yang berbeda. Menariknya, kemampuan musik pengguna tampaknya memiliki sedikit pengaruh pada kemampuannya untuk mengingat trek: pengguna dapat memilih sekitar 80% headset dalam kelompok yang diatur secara hierarki yang terdiri dari 30 orang atau lebih, terlepas dari asalnya.

4. Menjelaskan pengertian text, hypertext dan hypermedia

a. Pengenalan Tulisan Tangan (Handwriting Recognition)

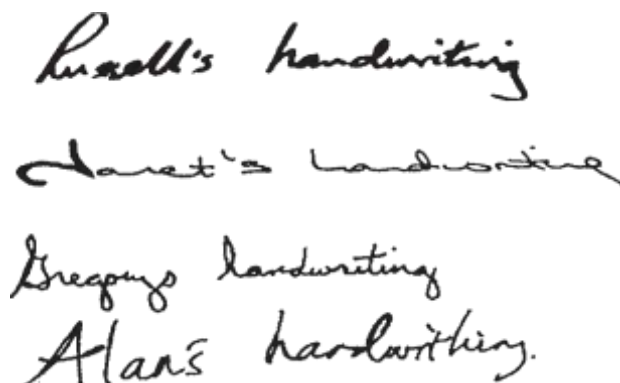
Saat kita berbicara, kita mengibaratkan tulisan tangan sebagai bentuk komunikasi yang amat natural. Ide untuk bisa memahami tulisan tangan amat menggoda, dan tulisan tangan sepertinya memberikan masukan tekstual dan grafis menggunakan alat yang sama. Namun, ada beberapa masalah dalam penggunaan tulisan tangan sebagai alat pengetikan dan kita akan membahasnya di bagian ini. Kami pertama-tama akan memeriksa mekanisme untuk mendapatkan informasi tulisan tangan, kemudian masalah interpretasi.

Teknik terpenting yang digunakan untuk memperoleh kaligrafi adalah digitizing. Denyut fluida diubah dengan pena menjadi serangkaian koordinat, dengan kecepatan satu setiap 1/50 detik (atau kira-kira kecepatan pemindaian konverter digital). Gerakan cepat menghasilkan poin dalam jarak jauh, dibandingkan dengan gerakan lambat: ini menciptakan kesalahan seketika dalam informasi, karena detail garis di antara titik-titik tersebut hilang, serta informasi tekanan.

Digitizing tablet telah ditingkatkan dengan masuknya layar tipis di bagian atas untuk menampilkan informasi dan menghasilkan kertas elektronik. Informasi yang direkam pada pemindai dapat dengan mudah dilihat kembali atau disimpan dan dilihat kembali untuk referensi di masa mendatang.

b. Tulisan Tangan (Handwriting Recognition)

Perbedaan antara tulisan tangan orang sangat besar (lihat gambar); Selain itu, tulisan tangan seseorang bervariasi dari hari ke hari dan berkembang dari tahun ke tahun.



Gambar 15.2 Tulisan tangan

Frasa umum yang setara juga umum dalam tulisan tangan, di mana huruf yang berbeda ditulis berbeda tergantung pada huruf sebelum dan sesudah. Sulit untuk melakukan ini dengan andal karena sistem pengenalan teks umum yang baik tidak berfungsi saat ini. Namun, ketika menulis surat secara terpisah, dengan sedikit pemisahan, keberhasilan sistem lebih dihargai, meskipun mereka harus dilatih untuk mengenali karakteristik pengguna yang berbeda.

c. **Text** (teks)

adalah format keluaran populer yang sangat digunakan didalam banyak kondisi. Semua teks tradisional ini memiliki sifat tertulis yang sama. Aristoteles mengatakan dalam buku puitisnya bahwa sejarah harus memiliki permulaan, pertengahan dan akhir, dan bahkan sejarah postmodern non-linier yang ditulis secara linier, walaupun kejadiannya tidak dapat dihubungkan secara kausal.

d. **Hypertext**

adalah untuk menghindari batasan teks ini dengan menyusunnya dalam kisi, bukan garis. Jika hypertext dirancang dengan baik, pengguna akan merasa lebih mudah untuk mengikuti idenya melalui jaringan daripada memaksakan arahan. Sistem hypertext biasanya berisi grafik, foto, dan media lain, selain teks. Sistem ini sering disebut sistem multimedia atau hypermedia, meskipun ketiga istilah tersebut digunakan secara bergantian. dokumen hypertext tidak hanya memulai progresi linier dan mengikutinya sampai akhir, tetapi berjalan ke arah yang berbeda, beberapa diakhiri sementara yang lain merujuk ke bagian dokumen yang berbeda.

e. **Hypermedia**

merupakan system hypertext yang mengintegrasikan tambahan media seperti sketsa, video dan bunyi.

Tujuan umum berfungsi untuk mendidikan:

- 1) kartun dan grafis
- 2) Struktur hypertextual

Masalah pada Hypermedia adalah:

- 1) Tersesat dalam hyperspace
- 2) Cakupan informasi yang tidak komplit
- 3) Sulit untuk mencetak dan membawa pergi

f. Animasi

Animasi adalah istilah untuk menambahkan gerakan ke gambar, membuatnya bergerak, berubah, dan berubah seiring waktu. Contoh animasi sederhana pada antarmuka adalah dalam bentuk jam. Alih-alih hanya membuat kursor utama selalu ditampilkan, banyak antarmuka sekarang menggunakan bitmap khas 16 x 16 yang membentuk kursor untuk menunjukkan informasi yang lebih kompleks.

g. Vidio dan Digital Vidio

Teknologi Compact Disc (CD) adalah sirkulasi didalam system multimedia: volume besar vidio, grafis, audio dan bacaan bisa disimpan dan ditemukan pada media yang relatif terjangkau serta mudah didapat. Berbagai metode dan metode kompresi memungkinkan Anda menulis lebih banyak data ke disk.

- 1) Compact Disc Interactive (CD-I)
- 2) eXtended Architecture (CD-XA)
- 3) Digital Video Interactive (DVI)/Universal Video Communications (UVC)

Contoh:

- 1) System berbasis DVI - Palanque
- 2) Sebuah prototipe dari system multimedia di mana pengguna mengunjungi situs web virtual. Gunakan vidio, gambar, teks, dan suara.
- 3) QuickTime Apple adalah standar untuk menyematkan vidio dalam sebuah antarmuka. Kompresi, penyimpanan, pemformatan serta sinkronisasi semuanya diatur sehingga berbagai aplikasi dapat menyatukan vidio dengan sempurna

h. Computer Vision

Computer vision atau juga disebut machine vision adalah ilmu pengetahuan yang mengembangkan teori-teori dan algoritma dimana informasi yang berguna mengenai dunia dapat secara otomatis diekstraksi dan dianalisis dari sebuah citra penelitian, sekumpulan citra, atau citra yang berurutan dari sebuah komputasi yang dibuat oleh komputer. Computer vision berhubungan dengan otomatisasi interpretasi citra untuk membuat berbagai pengukuran yang objektif atau untuk meningkatkan visibilitas ketelitian.

i. Gerakan (Gesture Recognition)

Gerakan adalah bagian dari Interaksi Manusia dan Komputer (IMK) yang telah menjadi pusat perhatian dalam sistem multimedia. Teknologi penangkap gerakan mahal, menggunakan computer vision atau lembar teknis khusus. Pengguna dapat menelusuri informasi melalui layar sentuh, joystick, atau suara.

5. Menjelaskan aplikasi multimedia sistem

Pengguna penyandang disabilitas menyangkut kepentingan istimewa yang ada kalanya dipenuhi system multimedia dan multimodal:

- a. Untuk tunanetra: pembaca layar, SonicFinder
- b. Fisik (disabilitas): input suara, pengenalan gerakan, dan system prediksi (contoh: keyboard interaktif)
- c. Kurang belajar (misalnya disleksia): berbicara dalam, dan luar.

a. Virtual Reality (VR)

Interaksi multimodal Multimedia memiliki batasan dan realitas virtual adalah simulasi komputer dunia tempat pengguna terbenam.

- 1) Headphone memungkinkan pengguna untuk 'melihat' dunia virtual
- 2) DataGlove Gesture Recognition (pelindung tangan kulit memiliki sensor partikel yang mengevaluasi posisi pada tangan dan jari.
- 3) Dengan Eyegaze, pemilik hanya bisa memutuskan arah menggunakan penglihatan mereka.

C. SOAL LATIHAN/TUGAS

1. Jelaskan mengenai system input, multimodal, dan multi media ?
2. Apa yang disebut dengan Speech dalam sebuah interface !
3. Jelaskan jenis-jenis non-speech sound !
4. Jelaskan mengenai text, hypertext, dan hypermedia ?
5. Jelaskan aplikasi multimedia sistem

D. DAFTAR PUSTAKA

Human-Computer Interaction, Third Edition, Prentice Hall, USA, 2003 by A.J. Dix, J.E. Finlay, G.D. Abowd and R. Beale

Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction: 5th IAPR TC 9 Workshop, MPRSS 2018, Beijing, China, August 20, 2018, Revised Selected Papers by Friedhelm Schwenker, Stefan Scherer

Multimodal Human Computer Interaction and Pervasive Services (Premier Reference Source) by Patrizia Grifoni

Multimodal Signal Processing: Theory and applications for human-computer interaction by Jean-Philippe Thiran, Ferran Marques, Herve Bourlard

The Human-Computer Interaction Handbook Fundamentals, Evolving Technologies and Emerging Applications, Second Edition (Human Factors and Ergonomics) by Andrew Sears, Julie A. Jacko.

Situs :

<https://b-ok.asia/dl/3265727/2c6902>

<https://b-ok.asia/book/5244285/6f1569>

<https://b-ok.asia/dl/853005/225be4>

<https://b-ok.asia/book/897778/d560e9>

<https://b-ok.cc/book/853211/86b4b6>

GLOSARIUM

Sensor input adalah untuk merasakan berbagai perbedaan yang terjadi di dunia nyata seperti temperatur, sentuhan, gaya, kelembapan, dan medan magnet.

Sistem Multimodal adalah sistem khusus yang mendukung komunikasi dengan pengguna melalui berbagai modalitas, seperti suara, tulisan tangan, gerak tubuh, emosi, dll.

Sistem Multimedia adalah media yang menggabungkan dua unsur atau lebih media media yang menggabungkan dua unsur atau lebih media.

Speech Recognition adalah sebuah kemampuan yang dimiliki oleh mesin atau aplikasi untuk mengidentifikasi kata dan frasa yang terdapat dalam bahasa lisan.

Speech Synthesis adalah sebuah kemampuan bicara manusia yang dibuat oleh manusia (artificial).

Text (teks) adalah format keluaran populer yang sangat digunakan didalam banyak kondisi.

Hypertext adalah untuk menghindari batasan teks ini dengan menyusunnya dalam kisi, bukan garis.

Hypermedia merupakan system hypertext yang mengintegrasikan tambahan media seperti sketsa, vidio dan bunyi.