

**VISION BASED HAND GESTURE  
RECOGNITION FOR HUMAN  
COMPUTER INTERACTION: A  
SURVEY**



## KESIMPULAN PADA TIAP PAPER

Siddharth S. Rautaray · Anupam Agrawal (2012) mengatakan bahwa “Analisis survei dan artikel yang luas menunjukkan bahwa teknik pengenalan isyarat tangan sering sensitif pada resolusi yang rendah, rata-rata framenya, kondisi pencahayaan, perubahan kondisi cuaca dan masalah umum pada sistem pengenalan isyarat tangan yang lainnya. Pada penelitian ini juga menyatakan bahwa penampilan berdasarkan gerakan tangan lebih disukai dari pada penampilan 3d pada sistem pengenalan isyarat tangan.”

Gandy M, Starner T, Auxier J, Ashbrook D (2000) mengatakan bahwa “ pada penelitian ini menunjukkan bahwa isyarat bisa digunakan dalam berbagai kondisi pencahayaan untuk mengontrol rumah secara otomatis. Melalui penggunaan varietas isyarat yang kontekstual, perangkat yang disambungkan pada konsep isyarat tersebut bisa di kontrol dan di batasi jumlah gerakan untuk mengontrolnya. Mereka juga telah menemukan bahwa perangkat ini dapat memberikan kenyamanan pada penduduk lansia dan juga bisa menyediakan fungsi tambahan sebagai diagnosa medis.”

Chen Q, Georganas ND, Petriu EM (2007) Mengatakan bahwa “ pada penelitian ini mereka mengusulkan pendekatan dengan dua tingkatan untuk mengenali gerakan tangan seperti real-time dengan kamera web tunggal sebagai perangkat input, tingkat ini bisa sangat memuaskan dengan akurasi yang sangat tinggi sekitar 90% . untuk tingkat tinggi pengenalan isyarat tangan mereka mengusulkan bebas konteks tata bahasa untuk menganalisa struktur sintaksis didasarkan pada terdeteksinya gerakan tersebut.”

Alon J, Athitsos V, Yuan Q, Sclaroff S (2009) Mengatakan bahwa “ mereka telah mengusulkan pendekatan gabungan, dimana mereka memperkirakan model gaussian untuk probabilitas dan observasi (seperti HMM), tetapi mereka juga menggabungkannya dengan transisi seragam model probabilitas DTW. Pendekatan ini telah digunakan untuk pengenalan karakter tulisan secara online.”

Bergh M, Gool L (2011) Mangatakan bahwa “ berdasarkan penelitian mereka, mereka mendeteksi bahwa 2d pada perobaan menampilkan gambar RGB berisi informasi yang cukup untuk mendeteksi gerakan dengan benar. Namun, eksperimen juga menunjukkan

kemungkinan untuk mengklasifikasikan hanya berdasarkan kedalaman informasi saja, yang berarti sistem dengan hanya ToF Kamera juga layak. Dalam tulisan ini, kedalaman Informasi diwakili sebagai 2D peta kedalaman, dan klasifikasi didasarkan pada 2D Haarlets.”

Nickel K, Stiefelhagen R (2003) Mengatakan bahwa “Pada percobaan ini mereka menggunakan model markov untuk melihat kinerja alat mereka dan mendeteksi gerakan rate (recall) dari 88% dan ketepatan 75%. Untuk 90% dari gerak tubuh, target menunjuk benar (salah satu dari delapan target) bisa diidentifikasi. Ketika melatih HMMs di subyek individu (orang-dependent), mereka melihat signifikan kinerja gain.”

B. Harrison, K. Fishkin, A. Gujar, C. Mochon, and R. Want (1998) mengatakan bahwa “dukungan lebih lanjut untuk melakukan pendekatan isyarat ini untuk para pengguna dengan antarmuka komputer datang dari karya Harrison et al, di mana para pengguna membuat gerakan di dunia nyata yang kemudian ditiru gerakan tersebut untuk mengontrol benda di dunia maya”

X. Huang, Y. Ariki, and M. A. Jack (1990) mengatakan bahwa “ buku ini berisi HMM berbasis sistem mendominasi pengenalan suara, dan keberhasilan dalam domain pidato telah melahirkan banyak usaha untuk memperpanjang metode HMM untuk pengenalan pola terkait bidang seperti pengakuan dokumen dan pengenalan tulisan tangan. Buku dikaji menjawab jelas perlu. Hal ini memperkenalkan sebagian teori dan teknik yang diperlukan untuk membangun lengkap HMM berbasis sistem pengenalan suara. Huang, Akira, dan Jack menggunakan pertama 90 halaman untuk menutupi metode umum dari pola pengakuan, pengolahan pidato-sinyal, dan pemodelan bahasa statistik.”

R. Orr and G. Abowd (2000) mengatakan bahwa “Untuk merasakan aktivitas pengenalan pola di rumah, upaya yang harus kita fokuskan berada pada cara menciptakan infrastruktur sensor dan komputasi dalam rumah, mulai dari lantai yang dapat mengidentifikasi orang-orang yang berjalan di atasnya”

S. Oviatt and P. Cohen (2000) mengatakan bahwa “ gesture pendant atau Gerakan liontin bisa kita gabungkan dengan pengenalan suara: Dalam pengaturan ini ucapan bisa digunakan

sebagai pengubah untuk satu penerapan yang lebih kecil dari yang ditetapkan dengan menggunakan gerakan.”

T. Starner, J. Weaver, and A. Pentland (1998) Mengatakan bahwa “Tujuan dari Pendant Gesture adalah untuk memungkinkan pemakainya untuk mengontrol elemen dalam rumah melalui gerakan tangan. Perangkat seperti peralatan hiburan yang ada di rumah dan pencahayaan ruang dapat dikontrol dengan gerakan sederhana.”

Cutler R, Turk M (1998) dan A. Corradini (2001) mengatakan bahwa “Fitur yang membedakan utama dari metode yang diusulkan yang ada bekerja di pengakuan isyarat adalah bahwa kita membutuhkan Metode yang tidak segmentasi spasial maupun temporal yang akan dilakukan sebagai preprocessing. Dalam banyak dinamis sistem pengakuan isyarat”

A. Quattoni, S. Wang, L.-P. Morency, M. Collins, and T. Darrell (2007) mengatakan bahwa “Hasil mereka menunjukkan bahwa HCRFs mengungguli kedua CRF dan HMMs untuk tugas-tugas pengakuan gerakan tertentu. Untuk gerakan lengan, yang Model HCRF multiclass melebihi HMMs dan CRF bahkan ketika Kisaran dependensi lama tidak digunakan, menunjukkan keuntungan dari belajar diskriminatif bersama. Untuk pengenalan obyek Data set, hasil kami menunjukkan bahwa menggabungkan dependensi antara variabel laten yang penting dan bahwa minimum formulasi pohon rentang dapat menjadi pendekatan yang baik untuk model yang lebih tinggi dihubungkan.”

Y. Zhu, G. Xu, and D. Kriegman (2002) mengatakan bahwa “Beberapa metode yang telah diusulkan dalam literatur untuk isyarat gerakan bercak yaitu, untuk mengenali gerakan dengan terus menerus sesuai dengan tidak adanya temporal yang diketahui segmentasinya”

H. Yoon, J. Soh, Y. Bae, and H. Yang (2001) mengatakan bahwa” pendeteksian tangan tetap menjadi tantangan tugas dalam banyak pengaturan kehidupan nyata. Masalah dapat disebabkan oleh berbagai faktor, seperti mengubah pencahayaan, rendahnya kualitas video dan gambaran yang tak jelas, resolusi rendah, sementara kemacetan atau kemunduran yang lain, dan kekacauan latar belakang. Biasanya digunakan visual yang isyaratnya bisa dideteksi dari tangan seperti warna kulit, tepi, gerak, dan latar belakang pengurangan

Y. Wu, T. S. Huang (2001) mengatakan bahwa” Untuk mencapai interaksi manusia-komputer secara alami untuk VE aplikasi, tangan manusia dapat dianggap sebagai alat input . gerakan tangan manusia adalah yang paling kuat untuk modal manusia dalam berkomunikasi. Namun, ekspresi dari gerakan tangan belum sepenuhnya dieksplorasi untuk aplikasi HCI. Dibandingkan dengan perangkat HCI tradisional, gerakan tangan kurang mengganggu dan lebih nyaman untuk menjelajahi tiga-dimensi (3D) di dunia maya.”

H. Zhou, T. S. Huang (2003) mengatakan bahwa” Teknik pengenalan isyarat tangan berbasis penglihatan dapat dibagi menjadi dua kategori: pendekatan berdasarkan penampilan- dan pendekatan berbasis model tangan 3D .”

P. Viola, M. Jones (2001) robust real-time mengatakan bahwa” Awalnya untuk tugas pelacakan wajah dan deteksi, mereka menggunakan pendekatan statistik untuk melacak dan mendeteksi berbagai macam wajah manusia.”

P. Viola, M. Jones (2001) Rapid object mengatakan bahwa” Dalam implementasi praktis, dalam aliran atensi digunakan untuk mempercepat kinerja pembelajaran algoritma. Pada tahap pertama dari proses pelatihan, ambang pengklasifikasi lemah disesuaikan cukup rendah sehingga 100% dari objek target dapat dideteksi sambil menjaga tingkat negatif palsu mendekati nol.”

Y. Freund, R. E. Schapire (1999) mengatakan bahwa” Dalam prakteknya ada fitur Haar-like tunggal yang dapat mendeteksi obyek dengan akurasi yang sangat tinggi.tapi, bisa saja dengan mudah untuk mencapai serangkaian pengklasifikasi lemah dengan akurasi yang sedikit lebih baik dari 50% menggunakan fitur Haar-like. The AdaBoost algoritma adalah pembelajaran metode untuk meningkatkan akurasi berdasarkan pada serangkaian tahap pengklasifikasi lemah tahap demi tahap.”

P. F. Felzenszwalb and D. P. Huttenloche (2006) mengatakan bahwa” menyarankan untuk melihat ke kedalaman, untuk membantu segmentasi tangan di latar depan dari objek dalam latar belakang. Pertama, kita dianggap stereo. Pada beberapa percobaannya dibuat menggunakan real-time visi stereo berdasarkan pekerjaan oleh Felzenszwalb et al.”

E.-J. Ong and R. Bowden (2004) mengatakan bahwa” mereka menguji detektor dengan database yang tak terlihat dari 2.509 gambar dan bentuk pengalififikasi dengan database yang tak terlihat dari 900 berlabel bentuk tangan. Dengan database ini, kami menemukan bahwa detektor dan pengalififikasi keduanya memiliki tingkat keberhasilan tiba-tiba tinggi 99,8% dan 97,4% masing-masing. Namun, kami mencatat bahwa gambar tangan untuk kedua pelatihan dan uji database memiliki cukup sederhana dan mirip latar belakangny.”

R. Reulke (2006) mengatakan bahwa” pada penelitian ini mereka meneliti sistem pencitraan dengan mengukur kemampuan jaraknya . Pendekatan untuk menggabungkan jarak dan Data gambar beresolusi tinggi disajikan. Ini adalah sebuah prasyarat untuk real time pencitraan 3D. Pada penelitian selanjutnya Akurasi pada prosedur perlu lebih lanjut diselidiki.”

Y.M. Kim, D. Chan, C. Theobalt, S. Thrun (2008) mengatakan bahwa” mereka telah menyajikan desain multi-view waktu-offlight sensor sistem pencatatan fusion. Menganalisa secara rinci dan mendalam ketidak akuratan pengukuran akan memungkinkan kita untuk mengimbangi sistematis kesalahan pengukuran TOF dan mengkalibrasi semua kedalaman dan data video ke dalam bingkai umum. Prosedur kalibrasi kami untuk kamera kedalaman sangat praktis dan mudah mengikatnya dengan prosedur kalibrasi kamera optik standar. Mulai dari data sekarang selaras, kita akan di masa depan menyelidiki cara-cara baru untuk meningkatkan bentuk dinamis dan rekonstruksi tekstur adegan dinamis acak.”

F. Wang and C. Zhang (2007) mengatakan bahwa” alat pemisah bergantung pada pengurangan dimensi berdasarkan rata-rata Lingkungan Margin Maksimalisasi (ANMM).”

T. Starner and A. Pentland (1994) mengatakan bahwa” model markov tersembunyi sudah digunakan selama bertahun-tahun terus menerus di bidang pengenalan ucapan atau suara , dan juga telah diterapkan dan berhasil untuk bidang pengenalan isyarat.”

D.A. Becker (1997) mengatakan bahwa” Seperti yang kita inginkan untuk mendeteksi dan menunjukkan gerakan on-line, merka telah memiliki urutan pengamatan untuk

menganalisis setiap kali bingkai baru telah diproses. Panjang BHE-fase bervariasi kuat dari satu gerakan ke yang lain. Oleh karena itu, kami mengklasifikasikan tidak hanya satu, melainkan serangkaian subsequences  $s_{1..n}$ , masing-masing mulai dari bingkai yang berbeda di masa lalu dan berakhir dengan frame  $t_0$  saat ini.”

J. Yang, W. Lu, and A. Waibel (1997) mengatakan bahwa” mereka telah mengusulkan sebuah model warna kulit statistik untuk melacak wajah manusia secara real-time. Mereka telah menunjukkan bahwa perbedaan penampilan warna kulit orang yang berbeda dapat dikurangi dengan normalisasi dengan analisis data. Menggunakan kebaikan-of-fit teknik, mereka yang telah diverifikasi lebih lanjut bahwa distribusi warna kulit masing-masing individu di bawah kondisi pencahayaan tertentu bisa ditandai dengan distribusi normal multivariat. Berdasarkan hasil ini, mereka memilikisulan model warna kulit adaptif untuk mengkarakterisasi manusia menghadapi kondisi pencahayaan yang berbeda. mereka telah menggunakan kombinasi linear dari parameter yang dikenal untuk memprediksi atau memperkirakan parameter baru tersebut.”

N. Jovic, B. Brumitt, B. Meyers, S. Harris, and T.Huang (1994) mengatakan bahwa” pada penelitiannya dia mendemonstrasikan penggunaan menunjukkan gerak tubuh untuk mencari benda-benda. Sistem mereka beroperasi pada berbagai fitur peta (intensitas, tepi, gerak, disparitas, warna). Jovic et a.”

L.R. Rabiner (1989) mengatakan bahwa” Untuk memperkirakan arah menunjuk, sangat penting untuk mendeteksi fase terus menerus secara tepat. Oleh karena itu, kita perlu tiga model fase terpisah: Tiga HMMs didedikasikan (MB, MH, ME) ditunjukkan secara eksklusif pada data milik fase mereka. Kami memilih topologi HMM yang sama (3 bagian, kiri-kanan) untuk masing-masing tiga model. Untuk masing-masing bagian, campuran 2 kepadatan gaussian merupakan probabilitas output. Mendapatkan nilai referensi untuk output dari model fase, kita melatih model nol  $M_0$  pada urutan fitur pendek (0.5sec) yang tidak termasuk sikap menunjuk. Untuk  $M_0$ , kami memilih HMM ergodic dengan 3 negara dan 2 gaussians per negara. Itu model dilatih dengan tangan berlabel BHE-fase menggunakan yang Baum-Welch persamaan estimasi ulang.”

