

USULAN TUGAS AKHIR

1 IDENTITAS PENGUSUL

Nama : Krizal Erdiyan Putra
NRP : 5109100021
Dosen Wali : Isye Ariesianti, S.Kom., M.Phil.

2 JUDUL TUGAS AKHIR

Rekonstruksi Objek 3D dari Citra 2D dengan Dekomposisi Area Geometrik dan Semantik menggunakan Fungsi Energi

Reconstruction of 3D Object from a 2D image with Decomposition Geometric and Semantic Region using Energy Function

3 URAIAN SINGKAT

Pada proposal ini akan diajukan suatu judul Tugas Akhir tentang metode yang secara otomatis membuat atau membangun model 3D dari suatu gambar foto tunggal 2D. Model yang akan dibangun memiliki kompleksitas dan kemiripan seperti buku ilustrasi *pop-up* anak-anak, dengan mencari daerah mana dalam foto yang menjadi objek daerah vertikal untuk di *pop-up* menjadi seperti terlihat muncul dari tanah. Topik tersebut diangkat dikarenakan merupakan topik yang unik dan masih jarang sehingga memiliki peluang penyelesaian masalah dengan metode baru yang bervariasi dengan kelebihan dan kekurangan masing-masing, dan dengan diwujudkannya judul Tugas Akhir ini maka akan menyumbangkan sebuah pemikiran atau ide baru tentang metode dalam area *scene understanding* yang perkembangannya sekarang sangat dibutuhkan khususnya dalam bidang *Computer Vision*.

Solusi yang akan diterapkan adalah dengan menggunakan metode dekomposisi citra ke dalam area citra yang konsisten secara *geometric* dan *semantic* dengan menggunakan *Energy function*. Tujuan utama dari proses ini adalah mendekomposisi atau membagi gambar ke dalam beberapa area dengan *semantic region labels* (yaitu, *road*, *sky*, *bus*, *car*, dst) dan *coherent geometric placement* (orientasi dan lokasi yang memperhatikan letak horizon). Label-label tersebut kemudian digunakan untuk mengetahui area mana dari gambar 2D yang berupa daerah yang akan di *pop-up* menjadi model objek 3D dengan metode “*Cut’n’Fold*”. Dengan memanfaatkan metode yang sebelumnya sudah ada, yaitu metode yang dikembangkan oleh Hoiem *et al.* Pada pengerjaan Tugas Akhir ini akan dilakukan suatu pendekatan metode yang berbeda dalam melakukan klasifikasi piksel citra. Sebelumnya selain itu juga terdapat pemaparan solusi lain untuk permasalahan ini yaitu metode yang diusulkan Saxena *et al.* dengan menggunakan model MRF dan Olga *et al.* dengan menggunakan CRF model.

Dengan pengerjaan Tugas Akhir ini, diharapkan dapat menyumbangkan suatu ide penyelesaian masalah dengan pendekatan metode yang baru untuk permasalahan *scene understanding* khususnya untuk masalah pemodelan 3D dari gambar 2D tunggal serta dengan metode ini nantinya bisa dikembangkan lagi.

4 PENDAHULUAN

4.1 LATAR BELAKANG

Perkembangan teknologi informasi pada bidang citra beberapa tahun terakhir ini telah menjangkau bentuk citra tidak hanya dalam 2D (Dimensi), namun sekarang sudah menjangkau teknologi untuk membangun sistem citra dalam bentuk 3D (Dimensi). Teknologi ini sangat banyak dipakai untuk berbagai bidang seperti pengembangan permainan 3D, sarana atau alat untuk membuat object 3D untuk keperluan desain, dan banyak bidang lainnya. Namun, untuk membuat sistem seperti itu sangatlah kompleks dan tidak mudah, seringkali membutuhkan alat-alat khusus, kemampuan khusus, dan menghabiskan waktu dalam prosesnya. Oleh karena itu banyak masyarakat umum yang enggan untuk memanfaatkan teknologi tersebut dan hanya mengandalkan orang yang ahli di bidang tersebut.

Selain itu, bidang fotografi sekarang juga sangat marak dan banyak diminati oleh masyarakat umum. Oleh karena itu kami berkeyakinan bahwa banyak orang yang akan menikmati suatu pengalaman untuk berjalan dan menyusuri secara nyata dari hasil foto mereka sendiri. Ketika seseorang melihat suatu gambar atau foto, yang mereka lihat bukanlah suatu bidang dengan pola warna dan tekstur, namun dunia dibalik gambar tersebut. Kami berkeinginan membuat suatu kemampuan yang sama kepada komputer, yaitu kemampuan mendapatkan model 3D dari sebuah gambar tunggal secara otomatis, istilah ini dalam dunia *Computer Vision* adalah *scene understanding*.

Ide tersebut terinspirasi dari model buku *pop-up* ilustrasi anak-anak, dimana objek vertical dari gambar muncul keluar dari *ground* (tanah). Model seperti inilah yang menjadi acuan dalam pembuatan tugas akhir ini, dimana gambar objek di dalam foto yang merupakan benda vertikal berdiri di atas tanah akan dibuat sebuah model 3D yang seolah-olah objek tersebut benar-benar muncul keluar dari *ground* (tanah) seperti model buku *pop-up* ilustrasi anak-anak. Dengan merujuk dari proyek penelitian-penelitian sebelumnya yang telah ada seperti metode yang telah dikembangkan oleh Hoiem *et al.*[2], Saxena *et al.*[3], & Olga *et al.*[7], dalam tugas akhir ini akan memodifikasi dan menggabungkan metode yang berbeda sehingga dapat menghasilkan hasil yang lebih baik dan metode yang lebih efisien dari metode yang telah diterapkan sebelumnya.

4.2 RUMUSAN MASALAH

Rumusan masalah yang diangkat dalam Tugas Akhir ini dapat dipaparkan sebagai berikut:

1. Bagaimana membangun sistem yang dapat mengkonversi gambar 2D tunggal menjadi gambar 3D secara otomatis.

2. Bagaimana merepresentasikan proyeksi tunggal 2D ke 3D yang secara matematika memiliki kemungkinan solusi dengan jumlah tak terbatas.
3. Bagaimana membangun metode untuk tipe gambar yang memiliki objek latar depan.
4. Bagaimana menerapkan metode untuk mendapatkan hasil yang akurat dan konsisten dalam proses klasifikasi piksel.
5. Bagaimana menerapkan dan menggabungkan rekonstruksi geometrik dan semantik untuk mendapatkan hasil 3D yang sesuai.
6. Bagaimana menerapkan *image segmentation* dengan menggabungkan beberapa variabel menggunakan *Energy function*.

4.3 BATASAN MASALAH

Permasalahan yang dibahas dalam Tugas Akhir ini memiliki beberapa batasan, diantaranya sebagai berikut:

1. Gambar atau foto yang dijadikan masukan 2D adalah gambar di luar ruangan (*outdoor*), bisa pemandangan alam (yaitu, gunung, bukit, pohon) dan buatan manusia (yaitu, gedung, mobil, manusia).
2. Diasumsikan gambar disusun dari sebuah bidang permukaan (*ground*) tunggal, objek planar yang mencuat dari permukaan (*ground*), dan langit.
3. Pemodelan metode dalam klasifikasi piksel menjadi beberapa *region* yang dibutuhkan dalam rekonstruksi 3D dengan menggunakan isyarat *geometric & semantic* serta dengan memanfaatkan *Energy Function* untuk mengkorelasikannya.

4.4 TUJUAN TUGAS AKHIR

Tugas Akhir ini memiliki tujuan yang rinciannya dapat dituliskan sebagai berikut:

1. Membangun sistem yang dapat mengkonversi gambar 2D tunggal menjadi representasi 3D.
2. Mendapatkan suatu metode baru yang berbeda dari hasil penelitian sebelumnya dalam hal klasifikasi piksel untuk segmentasi label gambar.
3. Memberikan pengalaman secara nyata kepada *user* dalam melihat dan menyusuri gambar secara 3 dimensi dari hasil gambar foto mereka sendiri.
4. Mendapatkan hasil yang lebih baik untuk gambar *input* yang memiliki objek latar depan.

5 TINJAUAN PUSTAKA

5.1 Piksel

Piksel adalah unsur gambar atau representasi sebuah titik terkecil dalam sebuah gambar grafis yang dihitung per inci. Piksel sendiri berasal dari akronim bahasa Inggris *Picture Element* yang disingkat menjadi *Pixel*. Pada ujung tertinggi skala resolusi, mesin cetak gambar berwarna dapat menghasilkan hasil cetak yang memiliki lebih dari 2.500 titik per inci dengan pilihan 16 juta warna lebih untuk setiap inci, dalam istilah komputer berarti gambar seluas satu inci persegi yang bisa ditampilkan pada tingkat resolusi tersebut sepadan dengan 150 juta bit informasi.

Monitor atau layar datar yang sering kita temui terdiri dari ribuan piksel yang terbagi dalam baris-baris dan kolom-kolom. Jumlah piksel yang terdapat dalam sebuah monitor dapat kita ketahui dari resolusinya. Resolusi maksimum yang disediakan oleh monitor adalah 1024x768, maka jumlah pixel yang ada dalam layar monitor tersebut adalah 786432 piksel. Semakin tinggi jumlah piksel yang tersedia dalam monitor, semakin tajam gambar yang mampu ditampilkan oleh monitor tersebut.

5.2 VRML

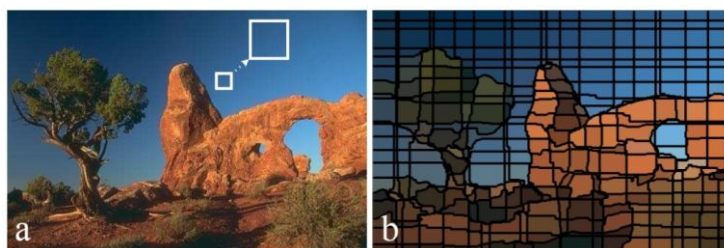
VRML (*Virtual Reality Modeling Language*) adalah sebuah bahasa pemrograman yang digunakan untuk mendeskripsikan gambar tiga dimensi (3D) dan suatu fasilitas yang memungkinkan pengguna bisa berinteraksi dengan objek 3D tersebut. Dengan menggunakan VRML, Pengguna dapat membangun sebuah rangkaian gambar visual ke dalam pengaturan web sehingga pengguna dapat berinteraksi dengan melihat, bergerak, berputar, dan sebaliknya berinteraksi dengan model 3-D. Sebagai contoh, Pengguna dapat melihat dan menggunakan kontrol untuk berpindah ruangan seperti yang dialami jika kita berjalan melalui ruang nyata. Untuk melihat file VRML, Anda memerlukan penampil VRML atau *browser*, yang dapat menjadi *plug-in* untuk *browser Web* yang sudah Anda miliki.

5.3 Segmentasi Citra

Dalam visi komputer, segmentasi citra adalah proses membagi citra digital menjadi beberapa segmen (set piksel, juga dikenal sebagai *superpixels*). Tujuan dari segmentasi adalah untuk menyederhanakan atau mengubah representasi gambar menjadi sesuatu yang lebih bermakna dan lebih mudah untuk dianalisis. segmentasi citra biasanya digunakan untuk menemukan obyek dan batas-batas (garis, kurva, dll) dalam gambar . Lebih tepatnya, segmentasi citra adalah proses untuk menempatkan label untuk setiap pixel dalam gambar sehingga pixel dengan label yang sama saling berbagi karakteristik visual tertentu.

Hasil dari segmentasi citra adalah seperangkat segmen yang secara kolektif mencakup seluruh gambar, atau satu set kontur yang diekstraksi dari gambar. Setiap piksel dalam suatu daerah yang sama memiliki hubungan dengan beberapa properti karakteristik, seperti warna, intensitas, atau tekstur [8].

Metode Segmentasi yang akan digunakan dalam Tugas Akhir ini adalah *Unsupervised Oversegmentation* dengan menggunakan algoritma *mean shift*. Dalam dunia *Computer Vision* dalam bidang segmentasi dikenal istilah *undersegmentation* yang memisahkan citra ke dalam sedikit kelompok area dan *oversegmentation* yang memisahkan citra ke dalam banyak kelompok area.



Gambar 1. *Unsupervised oversegmentation*

Contoh *Unsupervised Oversegmentation* dari sebuah citra menjadi *superpixel* dapat dilihat pada Gambar 1(b). Algoritma ini menunjukkan teknik yang meliputi topologi reguler sebagai properti dalam proses segmentasi. Gambar 1(a) menunjukkan contoh segmentasi citra yang bertujuan untuk memberi label semantik pada *pixel* atau *superpixel*. Permasalahannya adalah untuk menyelesaikan ambiguitas yang terjadi terhadap label semantik yang bisa saja berlabel langit, air, mobil, atau orang. Tetapi dalam konteks gambar keseluruhan yang diberikan kita dapat melihat kotak persegi tersebut berada di atas gunung, kotak persegi tersebut bukan merupakan orang, jalan, perahu, atau objek lainnya dan kami berkesimpulan kelas yang paling mendekati untuk kotak persegi tersebut adalah langit.

Oversegmentation yang dihasilkan, dibuat dengan menggunakan algoritma *Mean Shift* [9]. Segmentasi *Mean Shift* adalah teknik homogenisasi lokal yang sangat berguna untuk meredam perbedaan *shading* atau *tonality* pada obyek lokal. *Mean Shift* bekerja dengan menggantikan setiap pixel dengan rata-rata dari piksel di lingkungan kisaran- r dan yang nilainya dalam jarak d .

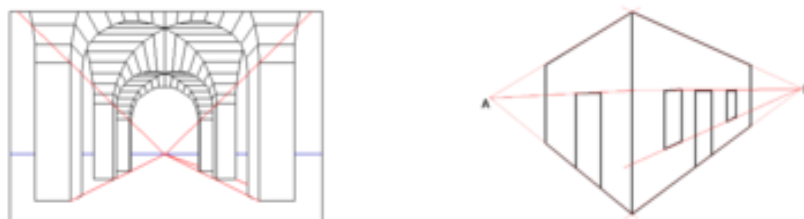
Fungsi *Mean Shift* biasanya membutuhkan 3 input:

1. Sebuah fungsi jarak untuk mengukur jarak antara piksel. Biasanya jarak *Euclidean*, tetapi dapat juga dengan menggunakan fungsi jarak lainnya seperti fungsi jarak *Manhattan*.
2. Radius. Semua piksel dalam radius ini (diukur menurut jarak di atas) akan diperhitungkan untuk perhitungan selanjutnya.
3. Perbedaan nilai. Dari semua piksel dalam radius r , akan diambil hanya mereka yang berada dalam nilai-nilai perbedaan ini untuk menghitung rata-rata.

5.4 Vanishing Point

Vanishing point merupakan satu dari beberapa kemungkinan titik dalam sebuah gambar 2D di mana garis-garis yang sejajar berkumpul pada satu titik sumber gambar 3D.

- *linear perspective* menggambar 1-3 *vanishing points*.
- *curvilinear perspective* menggambar dengan 4 atau 5 *vanishing points*, pada 5 titik perspective *vanishing points* dipetakan kedalam sebuah lingkaran dengan 4 *vanishing points* pada *cardinal headings* utara, barat, selatan, timur and satu di lingkaran asal.
- *reverse perspective* menggambar dengan *vanishing points* yang ditempatkan di luar lukisan dengan ilusi bahwa mereka "di depan" lukisan.



Gambar 2. *Vanishing Point*

Gambar 2 adalah contoh ilustrasi *gambar One-point perspective projection* (kiri) dan *two vanishing point* (kanan).

5.5 Referensi Aplikasi Sejenis

Beberapa referensi aplikasi sejenis yang sudah ada antara lain:

a. Make 3D

Metode yang diterapkan adalah dengan menggunakan *Markov Random Field* (MRF) untuk lokasi 3D dan orientasi terhadap area planar kecil pada citra. Hubungan antara fitur gambar dan lokasi / orientasi bidang didasarkan pada proses *learning*. Juga hubungan antara berbagai bagian dari gambar yang ditemukan menggunakan *Supervised Learning* (*Supervised Learning* adalah sebuah teknik *Machine Learning* untuk memprediksi nilai dari suatu fungsi yang diperoleh dari proses *training* data. Data *training* terdiri dari pasangan objek *input* dan *output* yang diinginkan). Petunjuk dasar di sini disebut "*superpixel*" yang pada gilirannya berhubungan dengan lokasi / orientasi bidang objek. Sebuah basis data yang terdiri dari berbagai jenis citra dibuat terlebih dahulu, dan kemudian peta kedalaman citra ini dapat dideteksi dengan peralatan yang mendukung. Bergantung kepada *likelihood* dari peta *superpixel* dan *scanner* peta kedalaman, maka dapat dengan mudah untuk membangun struktur 3D dari gambar 2D. Langkah perhitungan *superpixels* kemungkinan adalah memanggil *inferensi MAP* [3].

b. Automatic Photo popup

Dalam metode struktur 3D virtual dari gambar tunggal dibuat benar-benar otomatis. Hal tersebut nampak sebagai gambar diletakkan di tanah dan kemudian daerah-daerah yang harus tegak lurus dengan tanah yang "muncul" sebagai objek vertikal. Salah satu tujuan adalah untuk mencapai rekonstruksi 3D yang cepat.

Metode ini mengasumsikan bahwa gambar input adalah citra di luar ruangan bisa alam dan buatan manusia (bangunan) dan juga mengasumsikan bahwa satu citra terdiri dari objek tanah tunggal, dan *piece-wise* benda planar mencuat dari tanah di sudut kanan, dan ada juga langit dalam citra. Berdasarkan asumsi ini, aplikasi ini membangun sebuah model kasar 3D dari satu gambar dengan mengelompokkan setiap pixel menjadi tiga label yaitu "*Ground*", "*Vertikal*" atau "*Sky*" dan memperkirakan posisi cakrawala. Warna, tekstur, lokasi gambar, dan fitur geometris semua isyarat yang berguna untuk menentukan label ini dipergunakan [2].

c. Auto 3D

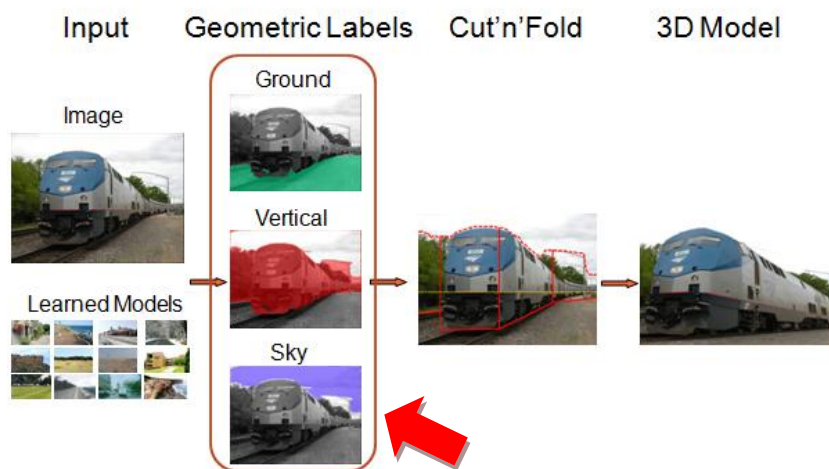
Tujuan dari metode ini adalah untuk membuat rekonstruksi 3D yang bagus dan menyenangkan dari sebuah gambar tunggal cepat dan otomatis. Metode membatasi ruang lingkup dari citra ke citra perkotaan, yang terutama terdiri dari bangunan-bangunan.

Pertama-tama diasumsikan sebuah citra memiliki sebuah *ground* dimana terletak pada bidang horizontal dan tembok vertikal dimana batas antara *ground* dan tembok merupakan garis *polyline* yang kontinu. Tujuan lain dari metode ini adalah untuk

meningkatkan efisiensi algoritma rekonstruksi 3D, dimana menjadi semakin dan semakin signifikan ketika resolusi dari citra tersebut tinggi. Metode yang digunakan adalah dengan menggunakan model *Conditional Random Field* (CRF) dengan menggabungkan variasi tipe yang berbeda (*appearance, geometric properties and context*) [7].

6 METODOLOGI

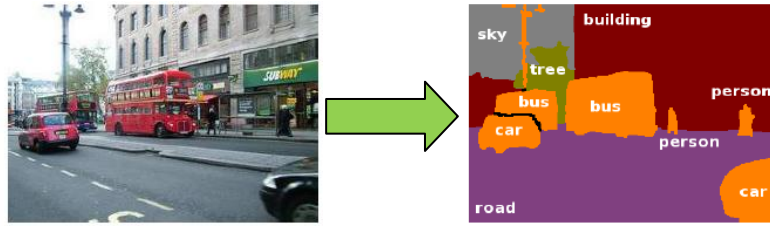
Tujuan utama dari pembuatan Tugas Akhir ini adalah membangun suatu sistem yang dapat merekonstruksi model 3D dari gambar 2D tunggal secara otomatis yang memiliki kesamaan karakter dengan buku *pop-up* ilustrasi anak-anak. Berdasarkan hasil riset yang sudah ada yang menjadi bahan rujukan utama kami, Hoiem *et al.*[2] yang memiliki algoritma estimasi model 3D seperti diagram pada Gambar 3. Gambar input 2D pertama-tama diklasifikasikan menjadi *geometric labels* menjadi 3 kategori “Ground”, “Vertical”, dan “Sky” dengan menggunakan *Learned Model*. Setelah itu berdasarkan kategori label tersebut, dilakukan pencocokan garis segmen (*fitted segment*) untuk daerah perbatasan antara label *ground-vertical* dengan bantuan pendekatan garis *horizon* yaitu garis kuning putus-putus. Kemudian dengan menggunakan teknik “Cut’n’Fold” yaitu dengan melipat (*fold*) daerah garis solid merah & memotong (*cut*) daerah garis merah putus-putus akan didapatkan daerah yang akan di *popup* menjadi model 3D, yang akan di proyeksikan ke dalam koordinat 3D pada VRML.



Gambar 3. Overview method

Namun dalam pengerjaan Tugas Akhir ini akan dilakukan pendekatan model yang berbeda dalam proses pemberian label untuk tiap-tiap piksel (*pixels classification*) dari gambar yang hanya menggunakan fitur *Geometric Labels* (daerah yang diberi kotak merah pada diagram Gambar 3). Metode sebelumnya menggunakan *superpixel* dan *multiple constellation* untuk segmentasi yang dilanjutkan dengan menggunakan algoritma *machine learning decision tree* berdasarkan fitur *Geometric Labels* saja yang memiliki beberapa kelemahan, diantaranya adalah terjadi ketidakkonsistensian dalam pemberian label dan menghasilkan hasil yang buruk untuk rekonstruksi 3D gambar yang memiliki objek latar depan (*foreground*).

Metode yang akan digunakan adalah metode yang diusulkan oleh Gould *et al.* [1] yaitu dengan menggunakan metode pendekomposisian gambar ke dalam area gambar yang konsisten secara *geometric* dan *semantic* dengan menggunakan *Energy function*.



Gambar 4. Segmentasi

Tujuan utama dari proses ini adalah mendekomposisi atau membagi gambar ke dalam beberapa area dengan *semantic region labels* (yaitu, *road, sky, bus, car, dst*) dan *coherent geometric placement* (orientasi dan lokasi yang memperhatikan letak horizon) dengan ilustrasi dengan Gambar 4, dengan menggabungkan kedua aspek tersebut kita bisa menghindari masalah ketidakkonsistensian (seperti jalan yang vertikal).

Diberikan sebuah citra I dan parameter model θ , nilai *Energy Function* yang menggabungkan keseluruhan descripsi dari gambar antara lain: asosiasi *pixel-to-region* \mathbf{R} , label kelas semantik \mathbf{S} , geometrik \mathbf{G} , kenampakan \mathbf{A} dan lokasi dari *horizon* v^{hz} . Persamaannya ditunjukkan oleh Persamaan (1).

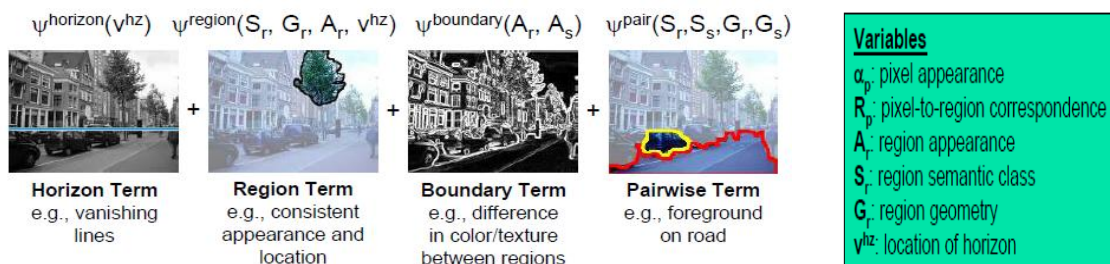
$$E(\mathbf{R}, \mathbf{S}, \mathbf{G}, \mathbf{A}, v^{hz}, \mathbf{K} \mid I, \theta) = \quad (1)$$

$$+ \psi^{\text{horizon}}(v^{hz}) \quad (2)$$

$$+ \psi^{\text{region}}(\mathbf{S}_r, \mathbf{G}_r, \mathbf{A}_r, v^{hz}) \quad (3)$$

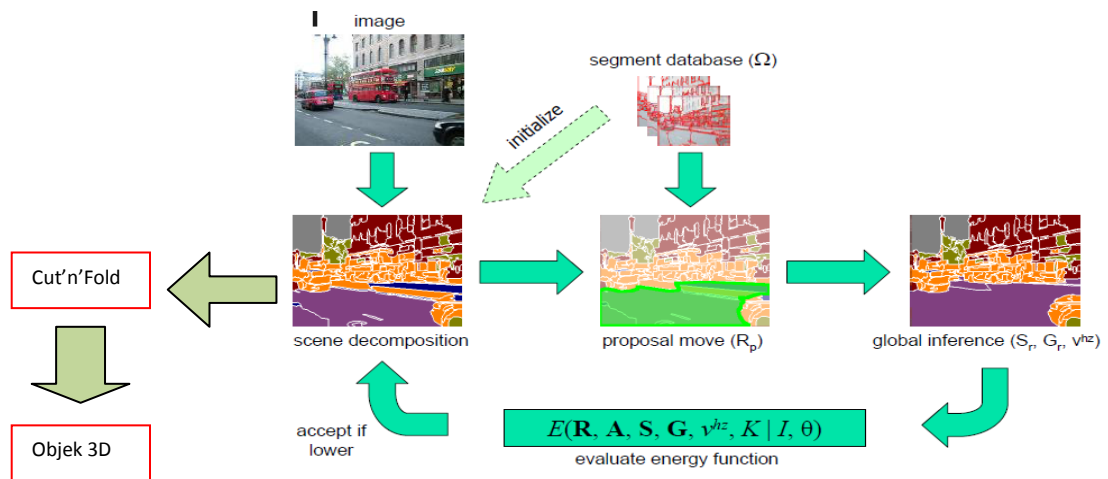
$$+ \psi^{\text{boundary}}(\mathbf{A}_r, \mathbf{A}_s) \quad (4)$$

$$+ \psi^{\text{pair}}(\mathbf{S}_r, \mathbf{S}_s, \mathbf{G}_r, \mathbf{G}_s) \quad (5)$$



Gambar 5. Variable Energy Function

Gambar 5 adalah diagram rumus *Energy Function*, dengan menggabungkan karakteristik dari *Horizon Term* dari Persamaan (2), *Region Term* dari Persamaan (3), *Boundary Term* dari Persamaan (4), dan *Pairwise Term* dari Persamaan (5). Dan Gambar 6 berikut adalah gambar diagram metode untuk dekomposisi citra.



Gambar 6. Alur *Decomposition*

Untuk membuat segmentasi dengan hasil yang baik dimulai dengan beberapa *over-segmentation* yang berbeda dari citra. Dengan menggunakan algoritma *mean-shift* [9] dengan tiga pengaturan parameter yang berbeda juga dengan memvariasikan parameter spasial dan jangkauan *bandwidth*, akan dihasilkan *over-segmentation dictionary* atau *segment database* (Ω). Kami kemudian menambahkan semua subset dibangun oleh proses ini (termasuk segmen awal) ke dalam *dictionary* atau *segment database* (Ω). Prosedur ini menghasilkan *proposal moves* (R_p) bervariasi. Tahap implementasi yang pertama adalah dengan melakukan *training data* terhadap 715 gambar daerah perkotaan dan pedesaan dari koleksi dataset gambar yang dipublikasikan: LabelMe [5], MSRC [11], PASCAL [6], dan Geometric Context (GC) [10].

Selanjutnya dengan menggunakan model yang telah didapatkan dari proses *training* sebelumnya, gambar 2D yang digunakan sebagai input diklasifikasi menjadi beberapa area label. Kemudian gambar dengan label area tersebut diolah dengan menggunakan metode “Cut’n’Fold” dan melakukan estimasi *horizon position* dan *Vanishing point* untuk membangun model 3D dengan menggunakan VRML.

7 JADWAL KEGIATAN TUGAS AKHIR

Tugas Akhir ini diharapkan dapat dikerjakan menurut jadwal sebagai berikut:

No.	Kegiatan	Bulan											
		Maret 2013			April 2013			Mei 2013			Juni 2013		
1.	Penyusunan Proposal Tugas Akhir												
2.	Studi Literatur												
3.	Analisa dan Perancangan												
4.	Implementasi												
5.	Pengujian dan Evaluasi												
6.	Penyusunan Buku Tugas Akhir												

8 DAFTAR PUSTAKA

- [1] S.Gould, R.Fulton, and D.Koller, "Decomposing a Scene into Geometric and Semantically Consistent Regions," *ICCV*, 2009.
- [2] D. Hoiem, A. A. Efros, and M. Hebert, "Automatic photo pop-up," *SIGGRAPH*, 2005.
- [3] A. Saxena, M. Sun, and A. Y. Ng, "Make3D: Learning 3-D scene structure from a single still image," *PAMI*, 2008.
- [4] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "Texton- Boost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation," *ECCV*, 2006.
- [5] B. C. Russell, A. B. Torralba, K. P. Murphy, and W. T. Freeman, "Labelme: A database and web-based tool for image annotation," *IJCV*, 2008.
- [6] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. (2007) The PASCAL Visual Object Classes Challenge 2007 (VOC2007). [Online]. <http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2007/>
- [7] Olga Barinova et al., "Fast Automatic Single-View 3-d Reconstruction of Urban Scenes," *ECCV*, 2008.
- [8] Wikipedia. Image Segmentation. [Online]. http://en.wikipedia.org/wiki/Image_segmentation
- [9] D.Comaniciu, P.Meer, and S.Member, "Mean shift: A robust approach toward feature space analysis," *PAMI*, 2002.
- [10] D. Hoiem, A. A. Efros, and M. Hebert, "Geometric context from a single image," *ICCV*, 2005.