

# Pengantar dan Survey Tentang Optical Music Recognition

Kevin Purwito

Program Studi Teknik Informatika, Universitas Multimedia Nusantara, Tangerang, Indonesia  
kpkevinp55@gmail.com

Diterima 11 April 2014

Disetujui 17 Juni 2014

**Abstract**—This paper describes about one of the many extension of Optical Character Recognition (OCR), that is Optical Music Recognition (OMR). OMR is used to recognize musical sheets into digital format, such as MIDI or MusicXML. There are many musical symbols that usually used in musical sheets and therefore needs to be recognized by OMR, such as staff; treble, bass, alto and tenor clef; sharp, flat and natural; beams, staccato, staccatissimo, dynamic, tenuto, marcato, stopped note, harmonic and fermata; notes; rests; ties and slurs; and also mordent and turn. OMR usually has four main processes, namely Preprocessing, Music Symbol Recognition, Musical Notation Reconstruction and Final Representation Construction. Each of those four main processes uses different methods and algorithms and each of those processes still needs further development and research. There are already many application that uses OMR to date, but none gives the perfect result. Therefore, besides the development and research for each OMR process, there is also a need to a development and research for combined recognizer, that combines the results from different OMR application to increase the final result's accuracy.

**Index Terms**—Music, optical character recognition, optical music recognition, musical symbol, image processing, combined recognizer

## I. PENDAHULUAN

Perkembangan ilmu pengetahuan dan teknologi yang pesat membawa dampak positif bagi kehidupan manusia. Dengan teknologi, pekerjaan manusia, di berbagai bidang, dapat menjadi lebih mudah, lebih baik dan lebih awet. Salah satu pengawetan yang bisa digunakan adalah dengan mengganti penggunaan kertas menjadi bentuk digital (*paperless*). Teknologi *paperless* kian populer karena keunggulan yang dimilikinya, antara lain lebih ramah lingkungan, murah, mudah disimpan, dicari, disunting, digandakan maupun disebarkan.

Upaya untuk mengubah dokumen kertas menjadi bentuk digital telah lama dilakukan. Teknologi untuk membaca tulisan sehingga dikenali oleh komputer, yaitu *Optical Character Recognition* (OCR), telah digunakan sejak tahun 1954 untuk sistem penyortiran alamat surat [1]. OCR juga digunakan untuk membaca

manuskrip-manuskrip lama, salah satunya yaitu manuskrip musik.

Pembacaan manuskrip musik ke dalam bentuk digital menggunakan suatu teknik pengembangan dari OCR, yang disebut dengan *Optical Music Recognition* (OMR). Berbeda dengan OCR yang membaca karakter, OMR yang dikembangkan sejak tahun 1960-an di Massachusetts Institute of Technology (MIT) [2], digunakan untuk secara khusus membaca notasi musik.

OMR terus dikembangkan untuk memudahkan pelestarian partitur-partitur musik, terutama partitur-partitur karya para komposer ternama, seperti Mozart, Beethoven dan Vivaldi. Setelah diubah ke dalam bentuk digital, maka musik akan bisa dipelajari dan juga dimainkan oleh generasi yang akan datang.

Paper ini terdiri atas lima bab. Bab pertama merupakan pendahuluan, diikuti oleh bab kedua yang berisi landasan teori yang mendukung penelitian ini. Bab ketiga merupakan alur kerja umum OMR. Selanjutnya, bab keempat berisi contoh implementasi OMR serta penelitian lebih lanjut pada OMR. Lalu bab kelima yang berisi simpulan paper ini.

## II. OPTICAL CHARACTER RECOGNITION, OPTICAL MUSIC RECOGNITION, DAN MUSICAL SYMBOL

### A. *Optical Character Recognition*

*Optical character recognition* merupakan teknologi pengenalan karakter dari masukan berupa gambar. Dengan kata lain, OCR merupakan suatu teknik *image processing* yang mampu menganalisis informasi grafis pada gambar yang diamati. Informasi grafis tersebut seperti bentuk garis, orientasi garis, posisi garis, jarak antar garis, dan sebagainya. Informasi grafis tersebutlah yang dianalisa sesuai dengan fitur-fitur yang dimiliki oleh suatu karakter.

OCR memiliki implementasi dan pengembangan yang berbeda, tergantung pada jenis karakter apa yang ingin dikenali. Salah satu pengembangan OCR adalah *Optical Music Recognition* (OMR) yang merupakan pengembangan OCR di bidang musik untuk mengenali partitur musik.

## B. Optical Music Recognition

*Optical Music Recognition* merupakan teknik pengenalan simbol musik dari gambar yang kemudian akan diubah ke dalam bentuk digital dan disimpan dalam bentuk MIDI (untuk dimainkan) atau MusicXML (untuk tampilan laman) [3].

MIDI adalah singkatan dari *Musical Instrument Digital Interface*, yaitu sebuah standar teknis yang digunakan untuk mendeskripsikan protokol, tampilan digital dan penghubung antar instrumen musik elektronik untuk saling berkomunikasi satu sama lain. MIDI membuat data digital musik menjadi dapat dimainkan oleh instrumen musik elektronik dan komputer.

MusicXML adalah sebuah standar penulisan XML (*eXtensible Markup Language*) untuk menyimpan data notasi musik dalam suatu partitur. Tujuan penggunaan standar penulisan XML ini adalah memudahkan representasi dan pembacaan notasi musik pada komputer.

Untuk dapat mengubah sebuah partitur musik menjadi bentuk data digital (MIDI atau MusicXML), OMR melakukan beberapa langkah pengerjaan, yaitu proses digitalisasi partitur menjadi *grayscale image*, *preprocessing*, *musical symbol recognition*, *musical notation reconstruction* dan *final representation construction* [4].

Setiap langkah pengerjaan tersebut terdiri dari proses yang kompleks dan menggunakan beberapa algoritma yang berbeda, dan akan dibahas lebih lanjut pada bab 3 paper ini.

## C. Musical Symbol

*Musical Symbol* adalah hasil usaha para musisi yang ingin mengekspresikan musik ke dalam simbol tertulis [5]. Notasi musik adalah visualisasi dari seluruh properti yang ada pada musik, seperti *pitch*, *dynamic*, waktu, dan *timbre* serta menunjukkan pilihan nada dan durasi.

Berikut ini merupakan uraian beberapa simbol yang umum digunakan pada notasi musik.

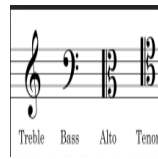
### 1. Staff



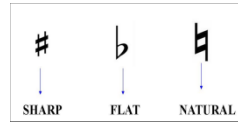
Kumpulan 5 *line* (garis) parallel dengan *space* (ruang) di antaranya. Di sinilah simbol musik lainnya ditulis.

### 2. Treble, Bass, Alto dan Tenor clef

Simbol musik pertama yang muncul pada setiap *staff* awal yang menunjukkan not jenis mana yang ada pada tiap *line* atau *space*.



### 3. Sharp, Flat dan Natural



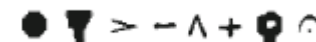
Simbol yang diletakkan di depan suatu not untuk menandakan perubahan *pitch*.

### 4. Beams



Simbol ini digunakan untuk menghubungkan beberapa not menjadi satu kelompok yang menunjukkan pembagian ritme.

### 5. Staccato, Staccatissimo, Dynamic, Tenuto, Marcato, Stopped note, Harmonic dan Fermata



Simbol-simbol ini menunjukkan penekanan khusus pada suatu ketukan.

### 6. Quarter, Half, Eight, Sixteenth, Thirty-second, dan Sixty-fourth notes



Simbol-simbol ini menunjukkan *pitch* dan durasi bunyi dari suatu not.

### 7. Quarter, Half, Eight, Sixteenth, Thirty-second, dan Sixty-fourth rests



Simbol-simbol ini menunjukkan durasi *silence* dalam suatu musik.

### 8. Ties dan Slurs



*Ties* digunakan untuk memanjangkan durasi dari

suatu not hingga ke ketukan berikutnya. *Slurs* digunakan untuk menandakan bahwa dua not harus dimainkan tanpa jeda diantaranya.

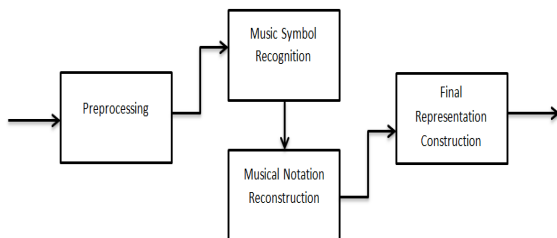
#### 9. *Mordent* dan *Turn*



Simbol-simbol yang mengubah pola *pitch* dari suatu not.

### III. ALUR KERJA OPTICAL MUSIC RECOGNITION

Secara garis besar, alur kerja OMR terdiri dari empat langkah yang ditunjukkan pada gambar 1.



Gambar 1. Alur kerja OMR

#### A. *Preprocessing*

Pada tahap ini, suatu partitur musik yang telah di-*scan* menjadi gambar digital akan dipersiapkan sebelum akan dikenali, yaitu dengan beberapa proses, antara lain:

##### 1. *Enhancement* dan *Noise Removal*

Proses ini memperbaiki dan mempertajam fitur-fitur gambar, seperti garis batas atau *contrast* untuk membuat gambar menjadi lebih jelas untuk dianalisa. Selain itu, proses ini juga melakukan pengurangan *noise*, penyaringan, magnifikasi dan lain-lain.

##### 2. *Binarization*

Proses ini mengklasifikasikan suatu gambar digital dengan format binari, yaitu angka 1 untuk merepresentasikan objek-objek (seperti simbol musik, *staff*, dan lain-lain) dan angka 0 untuk merepresentasikan latar belakang.

Tujuan pengklasifikasian gambar ini adalah untuk memisahkan antara bagian gambar mana yang bermanfaat untuk dianalisa dan mana yang tidak, sehingga akan mengurangi beban informasi yang perlu diproses oleh OMR.

Proses ini umumnya menggunakan metode Otsu [4]. Namun demikian, penelitian dan pengembangan lebih lanjut atas proses *binarization* masih diperlukan untuk memperoleh hasil yang lebih baik dan efisien.

#### B. *Music Symbol Recognition*

Pada tahap ini terjadi proses paling penting pada OMR, yaitu pengenalan simbol-simbol musik. Tahap ini terbagi menjadi beberapa bagian, yaitu:

##### 1. *Staff processing*

Pada bagian ini, OMR akan mengidentifikasi *staff lines* lalu akan menghapusnya, untuk mengurangi informasi yang tidak diperlukan.

##### 2. Pemrosesan simbol musik

Pada bagian ini, gambar yang sudah dihapus *staff line*-nya akan memperbaiki gambar, jika ada bagian gambar yang rusak lalu akan mengisolasi simbol primitif (*note*, *clef*, dan lain-lain) dan mengenalinya.

##### 3. *Classification*

Pada bagian ini, dilakukan ekstraksi fitur-fitur dari simbol yang dikenali dan juga terjadi pelatihan database. Ekstraksi fitur dan pelatihan database dilakukan untuk membuat OMR menjadi mampu mempelajari suatu simbol sehingga proses akan semakin baik dan cepat pada *recognition* berikutnya.

#### C. *Musical Notation Reconstruction*

Pada tahap ini, *staff line* yang sudah dihilangkan akan dikembalikan untuk digabungkan kembali dengan simbol-simbol primitif yang telah dikenali. Tahapan ini juga merupakan suatu tantangan sulit dalam OMR, mengingat diperlukannya suatu pemahaman sintaktik terhadap hubungan antar not pada suatu partitur. Ditambah lagi, musik bersifat dua dimensi, yaitu *pitch* secara vertikal dan waktu secara horizontal.

Oleh karena itu, digunakanlah suatu *grammar* yang digunakan sebagai dasar untuk menganalisa hubungan antar not, sesuai dengan yang terdapat pada *International Music Score Library Project (IMSLP)* [6]. *Grammar* tersebut berisi hubungan antara objek-objek grafis (*staff*, *note*, *rest*, dan lain-lain) dengan *parser*. *Parser* berisi komponen-komponen musik yang belum diberi label, yang kemudian akan di-*parse* dengan dua tahap yaitu pemberian label pada komponen-komponen tersebut, dan pendeteksian kesalahan.

Rossant dan Bloch [7] membagi *musical notation construction* menjadi dua bagian yaitu pendeteksian objek-objek menggunakan *low-level processing* dan pemberian hasil final berdasarkan informasi kontekstual dan aturan penulisan musik menggunakan *high-level processing*.

Pada tahap grafikal (*low-level processing*), dilakukan perhitungan derajat kecocokan antar objek (simbol musik), dengan beberapa aturan yaitu:

- *Accidental* dan *Notehead*: *accidental* diletakkan sebelum *notehead* dan selalu punya tinggi yang sama.
- *Notehead* dan *Dot*: *dot* selalu diletakkan di belakang

- atau di atas *notehead* pada jarak tertentu
- Di antara pasangan simbol lain tidak boleh saling bertabrakan.

Pada tahap sintaktik (*high-level processing*), dilakukan pengecekan aturan antara *tone*, *accidental* dan *meter*. Kumpulan simbol musik pada partitur akan dikenali sebagai kumpulan *accidental* setelah *clef* dan *meter* dari suatu partitur akan dihitung (jumlah ketukan di tiap baris).

#### D. Final Representation Construction

Pada tahap ini, partitur musik yang telah dikenali dan direkonstruksi akan disimpan ke dalam bentuk musik digital, antara lain MusicXML, MIDI, atau bisa pula disimpan dalam bentuk PDF.

Proses finalisasi ini umumnya menggunakan algoritma *Abductive Constraint Logic Programming* (ACLP) [8] yang menggabungkan algoritma *Abductive Logic Programming* (ALP) dan *Constraint Logic Programming* (CLP) menjadi sebuah framework. Framework ini mampu menghubungkan *high-level phase* yaitu tahap interpretasi simbol musik dengan *low-level phase* yaitu tahap pengenalan simbol musik, lalu menghasilkan keluaran berupa MIDI ataupun MusicXML.

### IV. IMPLEMENTASI DAN PENELITIAN SELANJUTNYA

Sudah ada banyak aplikasi yang mengimplementasikan *Optical Music Recognition*, antara lain Audiveris, SmartScore, SharpEye, PhotoScore, Capella-Scan, ScoreMaker, Vivaldi Scan, Gamera dan lain-lain [4].

Walaupun tahap-tahap yang dilakukan tiap aplikasi umumnya mirip, tetapi setiap aplikasi melakukan pendekatan dan menggunakan algoritma yang berbeda. Hal ini menyebabkan perbedaan hasil *recognition* yang dilakukan tiap aplikasi pada partitur yang sama. Selain itu, kualitas partitur yang dibaca juga sangat berpengaruh dan merupakan salah satu tantangan utama dalam pengembangan OMR [9].

Hal inilah yang mendorong E.P. Bugge dkk [10] untuk mengembangkan sebuah algoritma yang dapat membandingkan hasil dari beberapa aplikasi OMR kemudian mengombinasikannya sehingga menghasilkan *recognition* yang lebih akurat. Algoritma ini disebut *combined recognizer*.

Langkah-langkah dalam algoritma *combined recognizer* adalah (1) pertama-tama, lakukan OMR dengan masing-masing aplikasi secara independen; (2) lalu, ubah hasil OMR dari tiap aplikasi menjadi sebuah *sequence*; (3) gunakan *sequence alignment algorithm* untuk mencari kemiripan hasil antar aplikasi; (4) lakukan *voting* untuk memilih hasil yang paling sesuai; (5) kombinasikan hasilnya.

Dari percobaan E.P. Bugge dkk [10] tersebut,

terbukti bahwa hasil *recognition* dengan *combined recognizer* memiliki akurasi yang lebih tinggi. Oleh karena itu, selain pengembangan secara menyeluruh atas suatu aplikasi OMR yang lebih akurat, pengembangan algoritma *combined recognizer* juga merupakan suatu topik penelitian yang perlu dikembangkan untuk penyempurnaan OMR.

### V. SIMPULAN

*Optical Music Recognition* merupakan topik penelitian yang penting dan masih terus menerus dikembangkan untuk menambah akurasi dan efisiensi dari aplikasi-aplikasi yang sudah ada saat ini. Pada umumnya, OMR terdiri dari 4 tahap yaitu, *preprocessing*, *music symbol recognition*, *musical notation reconstruction*, dan *final representation construction* dimana tiap tahap memiliki proses-proses yang berbeda dan menggunakan algoritma yang berbeda pula. Hal ini juga yang membuat OMR menjadi kompleks dan memerlukan pengembangan dan perhatian secara menyeluruh.

Selain itu, diperlukan juga penelitian dan pengembangan algoritma untuk *combined recognizer* yang menggabungkan hasil beberapa aplikasi OMR yang sudah ada, karena setiap aplikasi memiliki kelebihan dan kekurangan masing-masing dalam mengenali partitur musik.

#### DAFTAR PUSTAKA

- [1] K. Danchisko, "Machines or Bust: Post Office Department Research and Development, 1945-1970", [diakses pada: 30 Maret 2014, 21.10], Alamat situs: [www.postalmuseum.si.edu/machinesorbust/p4.html](http://www.postalmuseum.si.edu/machinesorbust/p4.html)
- [2] D.H. Pruslin, "Automatic Recognition of Sheet Music", *Perspectives of New Music* Vol. 11, No. 1, Tenth Anniversary Issue (Autumn - Winter, 1972), pp. 250-254, 1966.
- [3] D. Bainbridge dan T. Bell, "An Extensible Optical Music Recognition", 19<sup>th</sup> Australian Computer Science Conference, 1996.
- [4] A. Rebelo, I. Fujinaga, F. Paszkiewicz, A.R.S. Marcal, C. Guedes, J.S. Cardoso, "Optical Music Recognition, State-of-the-Art and Open Issues", Springer-Verlag London Limited, 2012.
- [5] G. Read, "Music notation: a manual of modern practice", 2 edn. Taplinger, New York. ISBN: 0-8008-5459-4, 1969.
- [6] C. Raphael dan J. Wang, "New Approaches to Optical Music Recognition", 12<sup>th</sup> International Society for Music Information Retrieval Conference, 2011.
- [7] F. Rossant, I. Bloch, "Robust and adaptive OMR system including fuzzy modeling, fusion of musical rules, and possible error detection", *EURASIP J Appl Signal Process* 2007(1):160, 2007.
- [8] M. Ferrand, J.A. Leite, A. Cardoso, "Hypothetical reasoning: an application to optical music recognition", *Proceedings of the Appia-Gulp-Prode'99 joint conference on declarative programming*, pp 367-381, 1999.
- [9] D. Bainbridge dan T. Bell, "The Challenge of Optical Music Recognition", *Computers and the Humanities* 35: 95-121, 2001.
- [10] E.P. Bugge, K.L. Juncher, B.S. Mathiasen dan J.G. Simonsen, "Using Sequence Alignment and Voting to Improve Optical Music Recognition from Multiple Recognizers", 12<sup>th</sup> International Society for Music Information Retrieval Conference, 2011.