

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



**FEUP**

**Transcrição de Canto para Pauta  
Musical**

**PREPARAÇÃO DA DISSERTAÇÃO**

**Miguel Garcia**

Mestrado Integrado em Engenharia Electrotécnica e de Computadores

Orientador: Aníbal Ferreira (Prof.)

Junho de 2011



# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Breve introdução à escrita musical . . . . .	2
<b>2</b>	<b>Estado da Arte</b>	<b>7</b>
2.1	Conclusões da análise . . . . .	11
2.2	Contextualização e Objectivos . . . . .	13
<b>3</b>	<b>Ferramentas e Metodologias a adoptar</b>	<b>15</b>
3.1	Ferramentas de trabalho . . . . .	15
3.2	Metas e metodologia de trabalho . . . . .	15
<b>4</b>	<b>Resumo</b>	<b>17</b>
	<b>Referências</b>	<b>18</b>



# Abreviaturas e Símbolos

BPM	Batimentos por minuto
<i>f</i>	Forte
GUI	Graphical User Interface
MIDI	Musical Instrument User Interface
<i>p</i>	piano
SDK	Software Development Kit
WAV	Waveform Audio Format



# Capítulo 1

## Introdução

O presente documento serve de preparação para a dissertação no âmbito da transcrição de canto para a pauta musical. Está elaborado como documento final da unidade curricular de Preparação para a Dissertação promovendo uma familiarização com o tema da dissertação e uma investigação prévia sobre os assuntos abaixo mencionados.

O som é a percepção que temos ao receber no cérebro impulsos nervosos provenientes da parte periférica do sistema auditivo humano como consequência da chegada de uma onda sonora. É, de facto, a própria interpretação cerebral e não a onda sonora. É importante esclarecer esta diferenciação: a onda sonora é a vibração das partículas de um meio causando nós e ventres de pressão - uma onda mecânica, portanto - e apenas quando essa onda encontra o tímpano e tem algum efeito no cérebro é que se obtém o som (1). Apenas são audíveis ondas sonoras com frequências entre os 20Hz e os 20kHz, e a sua intensidade é igualmente um requisito à sua condição de perceptível. Um som pode ser musical ou não. O que define esta condição é, na maior parte dos casos, a relação entre os seus harmónicos. O harmónico de um som musical terá uma frequência múltipla inteira da frequência fundamental (2). A título de exemplo, um som produzido pela queda de um livro não será um som musical, ou seja, com uma nota definida, devido à relação variável dos seus harmónicos. A relação entre os harmónicos é dos mecanismos mais complexos da física do som pois define o conceito de timbre. Prova da dificuldade de compreensão do timbre é a própria definição da palavra pela American Standards Association: o atributo sensitivo que permite a distinção entre dois sons com a mesma intensidade e altura sonora (a altura sonora é o efeito psico-acústico causado pela variação da frequência); ou seja, a componente do som que não é nenhuma das que conseguimos efectivamente definir. Se quisermos fazer a analogia da audição à visão, o timbre seria a cor que difere os instrumentos musicais. Ao abordarmos a captação de trechos musicais por hardware temos de ter noção do funcionamento do sistema auditivo humano e de alguma física por trás do próprio som porque é isto que o sistema artificial terá de imitar. O microfone será análogo ao tímpano (contém, aliás, igualmente uma membrana) enquanto um processador cumprirá o papel da interpretação dada pelo cérebro. Mais à frente trataremos de forma mais dedicada o

software necessário, assim como um possível método de reconhecimento e de transcrição.

É vulgar considerar-se a voz cantada como o mais prático instrumento musical por ser parte integrante do ser humano. Estando nós tão familiarizados com tal aparelho pressupõe-se, portanto, que o seu domínio é geral e que a sua utilização é a mais intuitiva. Não desfazendo esta ideia, nem sempre é assim, e aqui a tecnologia pode ser também um grande auxílio ao permitir ao cantor aperceber-se mais claramente do que está a entoar e se o está a fazer correctamente. Por outro, lado pode igualmente aproveitar-se essa habitação ao trabalho com a voz para criar novos processos de escrita musical em que a parte mais trabalhosa se esboça com a voz e não com a mão, pacientemente, num papel.

## 1.1 Breve introdução à escrita musical

Todo o processo de ensino musical do sistema moderno se baseia na partitura musical que para além da pauta - um simples conjunto de 5 linhas e 4 espaços que serve de base à notação musical - contém todas as notas, pausas e muitos outros símbolos de significado musical relevante. É neste preceito que se apoiam os actuais músicos seja para disponibilizar a sua música de forma pessoal ou comercial, para aprender ou para ensinar. Este registo, apesar de o considerarmos quase ideal e bastante exacto, não deixa de ser um tipo de escrita que exige alguma formação musical e que é, para o copista menos experiente, um processo moroso. A música faz-se de sons e de silêncios representados por figuras musicais que indicam som durante um determinado intervalo de tempo e pausas associadas às figuras com valor análogo em silêncio. As figuras musicais são as 7 que se ilustram na figura 1.1















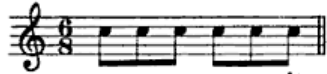

	Nota	Pausa	Nome
4			Semibreve
2			Mínima
1			Semínima
1/2			Colcheia
1/4			Semicolcheia
1/8			Fusa
1/16			Semifusa

Figura 1.1: Figuras e pausas musicais e tempo respectivo



Estas representações gráficas referem-se ao ritmo que, por sua vez, não é mais do que uma representação da duração temporal das notas com uma agregação própria. Um dos aspectos relevantes para entender a transcrição que será abordada nesta dissertação é a divisão em compassos, da qual dependerá a tal agregação das figuras. Esta escolha do tipo de compasso e a divisão musical a ela associada depende somente da interpretação a dar à linha melódica em questão porque apenas condicionará os tempos fortes da mesma, ou seja, a sua acentuação. Diz-se então que a métrica é um processo de extracção de informação ligada ao ritmo e à intensidade por parte do cérebro. Curiosamente, a escrita será totalmente dependente desta decisão mas a captação totalmente independente. Quando se pretende dividir a linha melódica em trechos de igual duração é necessário estabelecer uma base que será de 1 tempo ou 1,5 consoante o compasso seja simples ou composto, respectivamente. A figura que corresponde a 1 tempo é a mínima e para obter a base de 1,5 tempos usa-se a mínima pontuada como figura base. Há também os compassos binários e os ternários. Apesar de se poderem categorizar os compassos em mais tipos como quaternários, quinários e septanários, simplificaremos para as 2 tipologias mais abrangentes. O quaternário, por exemplo, passará a binário porque possui um número par de figuras base. Na figura 1.2 exemplificam-se alguns tipos de compasso.

**Compasso Binário Composto**   
cada colcheia representa  $\frac{1}{3}$  de tempo

**Compasso Ternário Composto**   
cada semicolcheia representa  $\frac{1}{3}$  de tempo

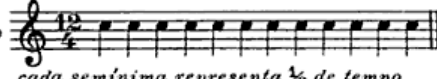
**Compasso Quaternário Composto**   
cada semínima representa  $\frac{1}{3}$  de tempo

Figura 1.2: Exemplo de tipos de compassos

A caracterização desta divisão irá figurar no início da partitura através de uma fracção. Se o numerador for 6, 9 ou 12 o compasso é composto e se for 2, 3 ou 4 é simples. É possível criar um compasso composto multiplicando o compasso simples pela fracção  $\frac{3}{2}$ . A figura 2.2 demonstra a transformação de um compasso simples no seu compasso composto correspondente.

<b>Binário</b>	<b>Ternário</b>	<b>Quaternário</b>
$2 \times 3 = 6$	$3 \times 3 = 9$	$4 \times 3 = 12$
$4 \times 2 = 8$	$4 \times 2 = 8$	$4 \times 2 = 8$

Figura 1.3: Exemplo da transformação de compassos

Se o numerador for múltiplo de 2, 4, 8 ou 16 o compasso é binário. No caso de ser 3 ou 9 é

considerado ternário. Se for 6, 12 ou 24 vai depender pois pode ser considerado binário ou ternário, consoante seja composto ou simples. O ideal será deixar o intérprete escolher a forma como quer escrita a sua linha melódica segundo o método que melhor reflecte o que este reproduziu. Em suma, a fracção do compasso define o seu valor indicando uma figura base com o denominador segundo a figura 1.4 e o numerador indica quantas figuras dessas deverão existir no compasso. Multiplicando o número de figuras pelo seu valor indicado na Figura 1.4 obtém-se o valor total de tempos no intervalo do compasso.





						
1	2	4	8	16	32	64

Figura 1.4: Valor relativo das figuras musicais

Olhemos ainda para o aspecto melódico. A divisão das frequências do espectro audível em notas chama-se escala musical. Escala musical temperada, aliás, uma vez que todos estes intervalos são iguais. Falamos da escala moderna porque nem sempre os intervalos foram iguais, privilegiando por vezes alguns intervalos entre frequências que seriam especialmente apelativos ao cérebro humano como a quinta perfeita ou a quarta perfeita. Actualmente, o intervalo entre notas divide-se vulgarmente em tons, sendo o intervalo entre 2 notas consecutivas chamado meio-tom (ou semi-tom). Uma vez que a oitava é o dobro da primeira frequência e antes de alcançar a oitava se passa por 12 notas, é trivial concluir que há 12 meios-tons entre cada oitava. É possível também ouvir falar em comas ou em cents, 2 medidas alternativas mas que neste caso não serão abordadas. A título de curiosidade, 1 meio-tom terá aproximadamente 100 cents e o coma obtém-se dividindo um tom em 9 intervalos. Temos então as 7 notas cujos nomes nos são normalmente familiares (dó, ré, mi, fá, sol, lá e si) e outras entre elas para as quais necessitamos de algo a que chamamos acidentes: o sustenido ( $\sharp$ ) e o bemol ( $b$ ). Estes símbolos aumentam e diminuem respectivamente meio-tom à nota na qual são aplicados. A escala é um conjunto de notas que podem ser utilizadas numa determinada tonalidade, variando consoante o tipo de escala (maior, menor, cigana, húngara, de blues, e muitas outras). É constituída por 7 notas (salvo raras excepções como a escala pentatónica) com intervalos entre elas que se mantêm constantes independentemente da nota inicial (que dá nome à escala), como se de uma forma se tratasse. Para identificar a escala basta identificar a relação entre as suas notas indicando os sustenidos e bemois que esta deve ter, junto à clave naquilo que se chama armação de clave e que é parte essencial da partitura sem a qual a transcrição nunca estaria completa. Um exemplo da armação de clave está representado na figura 1.5. É essencial compreender que a música será dramaticamente diferente se houver um erro na armação de clave.

Após esta análise, é então bastante fácil encontrar vantagens num software que permita a transcrição da voz cantada ou de outro instrumento musical monofónico para uma partitura musical. Não só permitiria uma função de “scoring” como um avanço no ensino da música como também

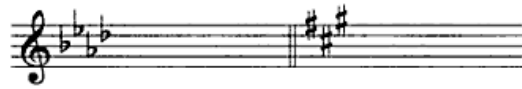


Figura 1.5: Exemplo de uma armação de clave em Lá♭ maior anulada no compasso seguinte

facilitaria a escrita musical tornando-a mais acessível a todos os que quisessem compor bastando para isso entoar a melodia desejada.



## Capítulo 2

# Estado da Arte

Actualmente existem alguns programas que reconhecem a nota captada quase imediatamente ou previamente gravada por exemplo num ficheiro de som. A robustez desse reconhecimento é muito variável devido à complexidade do processo que está por trás desta funcionalidade. Os programas mais promissores neste campo são o Music Masterworks e o Sing and See. Faz-se, neste ponto, uma crítica ao software analisado. Esta crítica teve de ser feita individualmente uma vez que não há bibliografia que ajude. Existe, efectivamente, uma panóplia de documentos académicos que abrangem a área das técnicas de processamento de sinal que estão por trás desses programas (3) e outros tantos sobre os programas de edição de escrita musical mas nenhum que faça uma comparação detalhada entre eles (4). Após a análise que será sintetizada na tabela 2.3, conclui-se - numa crítica pessoal - que nenhum dos programas apresentados tem a precisão de reconhecimento necessária para uma utilização eficiente do ponto de vista da produtividade musical. Apesar do seu interesse inovador no campo da utilização de tecnologias na música, não está ainda no ponto em que permite uma utilização geral e despreocupada por qualquer utilizador. É de sublinhar que apenas uma pequena parte do software experimentado realizava algum tipo de transcrição musical e nunca de forma totalmente satisfatória, mostrando que neste campo ainda é necessária a implementação de um método que satisfaça por inteiro as expectativas dos utilizadores. Através da tabela 2.3 é possível dividir os programas analisados em 2 tipos: programas de escrita musical que permitem escrever com facilidade partituras ainda que sejam complexas desenhando todos os símbolos musicais na pauta respectiva com possibilidade de reproduzir o resultado e os programas de gravação que permitem exportação em MIDI. Como o primeiro tipo de programa permite a leitura de MIDI, pode obter-se a transcrição musical de uma linha melódica gravando-a neste formato, usando para tal o outro tipo de software analisado. No entanto, falta uma integração entre as 2 técnicas que optimize o processo e o facilite para o utilizador. Não só permitiria efectuar a transcrição com a utilização de 1 só programa em vez de 2 mas também permitiria a correcção e adaptação da linha melódica captada à transcrição desejada. Explorando ambos os tipos de programa, vemos que aqueles que se dedicam à edição e criação de escrita musical se dedicam apenas à escrita

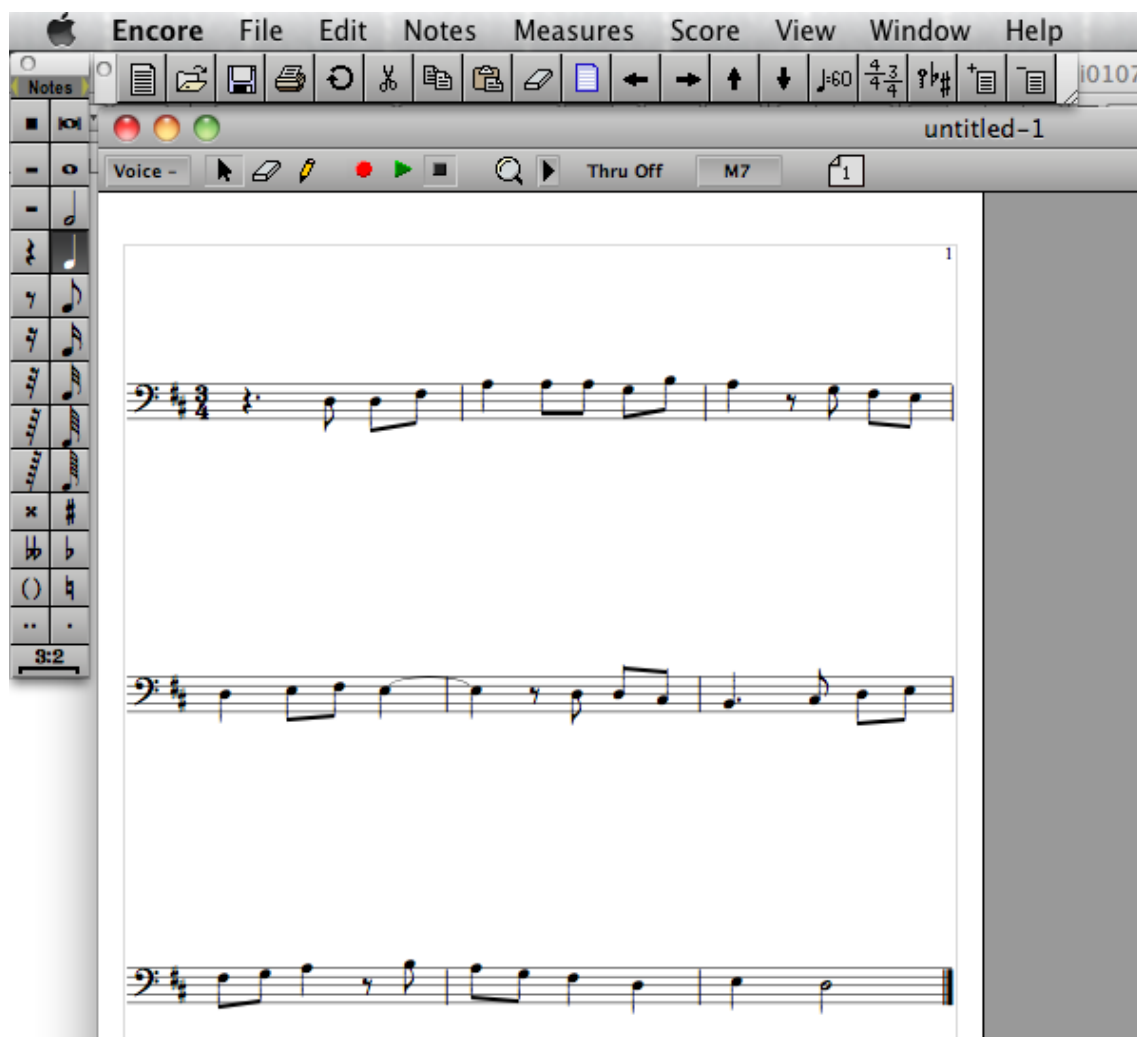


Figura 2.1: Exemplo do software Encore da Gvox

de complexas partituras com recurso à reprodução imediata. São baseados maioritariamente em MIDI, permitindo a associação de diferentes instrumentos musicais para cada pauta e facilitando assim a reprodução de partituras com múltiplas vozes. Precisam de ser altamente detalhados no que toca aos símbolos musicais e imitar com a maior naturalidade possível a musicalidade humana em indicações dinâmicas (como *accelarando*, *ritenutto*, entre outros).

No exemplo do programa Encore ilustrado na figura 2.1, vemos na barra lateral as figuras musicais a aplicar por arrastamento na pauta, podendo adicionar acidentes (bemóis, sustenidos ou bequardos) e incluindo outros símbolos que aparecerão na mesma janela quando seleccionados. Entre estes temos as claves, indicações de dinâmica, de ornamentação, indicações próprias de cada instrumento ou indicações de repetição. Pode ainda haver opções mais livres como a adição de cores, desenho livre ou campo de texto para que o criador possa exprimir-se da forma que desejar, ainda que não seja a mais normal. Na barra superior é visível um comando de *play* que permite a reprodução da partitura criada e um outro comando de gravação possibilitando a

captação directamente de um instrumento MIDI ligado ao computador onde corre o programa. Programas como o Encore são ferramentas essenciais para o trabalho de um músico compositor, sendo este o mais básico. Infelizmente, cada um destes programas tem um formato próprio que exige uma pluralidade de tipos de exportação quando se deseja criar um software que produza resultados interpretáveis por cada um deles. Este é apenas um exemplo elementar desta gama de programas mas foi o escolhido para representar a sua categoria devido à sua simplicidade.

Há ainda um segundo tipo de programa para o qual o Transcribe será representativo na figura 2.2.

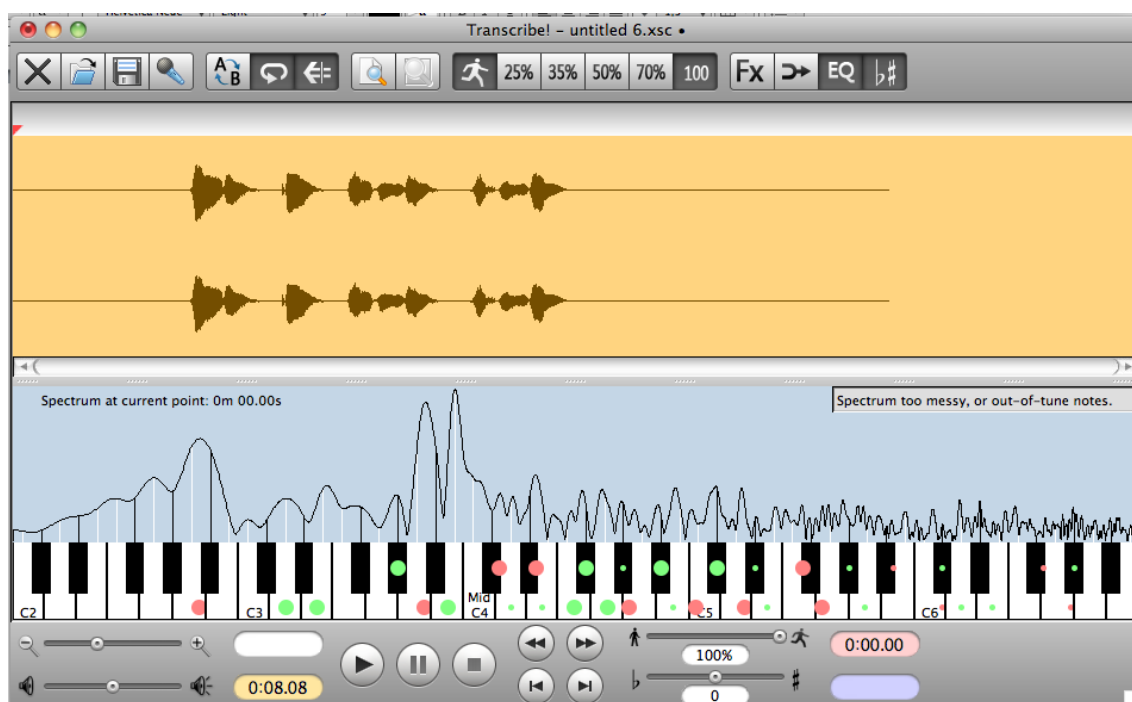


Figura 2.2: Exemplo do software Transcribe

Nesta categoria inscrevem-se os programas que captam linhas melódicas e reconhecem as notas captadas. Um dos métodos de reconhecimento é concentrar-se na frequência fundamental e tentar aproximar-se da frequência de uma nota conhecida. Este reconhecimento individual incorpora igualmente uma duração temporal e é concatenado ao longo de toda a gravação sendo normalmente mostrado num esquema de barras ou linhas horizontais, com possibilidade de se mostrar as respectivas notas num teclado de piano simulado na janela de forma a facilitar a sua visualização. Este processo tem uma sensibilidade variável de programa para programa (nunca ajustável dentro do mesmo software). Após o reconhecimento, podemos observar na janela do Transcribe a possibilidade de utilizar um botão deslizante que permite deslocar todas as notas para frequências superiores ou inferiores, assim como outra para o tempo. Estas funcionalidades não são, no entanto, comuns neste tipo de software. Tenta-se de seguida esclarecer os principais pontos de cada programa na tabela 2.3.

	Grava	Transforma MIDI	Mostra Partitura	Partitura Editável	Mostra Teclado	Scoring	Transforma de wav	Observações
<b>Sing&amp;See</b>	X	X	-	-	X	X	X	
<b>Encore</b>	X	-	X	X	X	-	-	
<b>Music MasterWorks</b>	N/D	X	X	X	X	N/D	N/D	<b>Trial</b>
<b>Singing Studio</b>	X	X	-	-	X	-	X	
<b>Singstar</b>	X	-	-	-	-	X	-	jogo disponível para consolas de jogos da <b>sony</b>
<b>Transcribe</b>	X	X	-	-	X	-	X	
<b>Twelve Keys</b>	-	X	-	-	X	-	X	Usa " <b>wave-pad sound editor</b> " para gravar
<b>Muse Score</b>	-	X	X	X	X	-	-	
<b>Carry a tune*</b>	X	-	X	X	X	X	-	
<b>Slowmp3</b>	-	X	-	-	X	-	X	
	N/D - Funções não disponíveis em <b>Trial</b>							
	* Análise com base em críticas literárias							

Figura 2.3: Tabela comparativa do Software analisado

É lógico inferir que a transcrição musical está intimamente dependente do reconhecimento melódico e que para otimizar a transcrição temos de adaptar desde logo o primeiro processo. A música escrita tem como base a matemática. Inicialmente dividem-se os sons em intervalos constantes de igual duração aos quais se chamam compassos, definição já explicada na introdução. É nestes intervalos que serão incluídos os sons captados, divididos em figuras diferentes consoante a sua duração (sempre relativamente ao valor base). A adaptação da captação deve começar desde logo pela inclusão de um metrónomo que condicione a reprodução musical ao tempo desejado. O valor do metrónomo deve ser indicado em bpm (batimentos por minuto) em caixa de texto designada para o efeito ou através de uma barra deslizante e deverá manter-se constante ao longo de toda a captação. Os batimentos que este reproduz devem ser compreendidos de forma perfeita pelo utilizador mas não podem de forma alguma afectar a gravação. Assim, propõe-se que sejam reproduzidos com a utilização de uns auscultadores, possivelmente com um filtro desse sinal na captação para o caso de este ser reproduzido pelas colunas do aparelho e com efeitos visuais (um contraste oscilante na janela do programa ao ritmo do metrónomo será uma possível solução). Após a transcrição da linha melódica seleccionada terá necessariamente de haver espaço



para alguma alteração manual, nomeadamente a nível de escolha de claves e tipos de compasso e correcção de notas às quais o processo automático não tenha sido sensível. Isto implica funções de adição e remoção de figuras. O ideal seria juntar a este programa uma funcionalidade de scoring que permitiria ao utilizador confirmar se o trecho musical que acabou de produzir corresponde de facto ao que deveria ter feito e, em caso negativo, onde errou. Uma vez mais num cenário ideal, poderia inclusivamente ser criado um conceito de escrita musical polifónica em tempo real, permitindo a gravação e processamento (incluindo transcrição) individual de faixas separadas - cada uma proveniente de um microfone - dando opção de as aglutinar formando uma só partitura. Esta funcionalidade permitiria a pequenos grupos instrumentais ou coros de câmara criar a sua própria música e escreve-la sem qualquer esforço, necessitando somente de um microfone para cada instrumento monofónico ou para cada voz.

## 2.1 Conclusões da análise

No programa utilizado - Singing Studio - foram então analisados algumas questões que teriam de ser modificadas para permitir a aplicação da transcrição para pauta musical e que serão enumeradas seguidamente. Antes de mais define-se um objectivo: pretende-se tornar o software utilizado numa plataforma que consiga gravar uma melodia monofónica, que a reconheça melódica e ritmicamente e que escreva a sua partitura, permitindo alguma edição no momento. Tendo este marco como ponto de chegada podemos dividir os problemas de aquisição e transcrição em dois tipos: problemas rítmicos e melódicos. As soluções passarão igualmente por dois processos: automático ou manual. O processo automático é preferível para os problemas sistemáticos e torna a utilização do programa mais simples. No entanto é menos fiável que a alternativa manual por obrigar à implementação de um conjunto de medidas que o utilizador pode não querer inteiramente. A solução manual é a ideal para quem quer um resultado mais preciso. Apesar de exigir que o utilizador perceba do que está a fazer, permite que os resultados sejam utilizáveis no futuro com muito mais confiança. Preferir-se-á a solução automática devido à vasta gama de produtos de software de extrema qualidade que permitem a alteração (e criação) de partituras, sendo apenas necessária a exportação no formato adequado.

Analise-se primeiramente a situação de uma gravação feita com as notas correctas na duração correcta mas na qual o utilizador não se apercebe do tempo a utilizar o que irá condicionar toda a transcrição. É apresentado um metrónomo durante a gravação que terá componente audio (terá de ser filtrada no processamento) e com componente visual (o ecrã poderá mostrar uma luz branca no centro, possivelmente todo o ecrã, que acende e apaga à velocidade do metrónomo). Haverá possibilidade de usar somente a componente visual do metrónomo impedindo a captação deste som pelo microfone, opção especialmente vantajosa na ausência de auscultadores. A velocidade será controlada por uma barra deslizante havendo igualmente a possibilidade de indicar por escrito qual será a velocidade desejada. Note-se que a base desta velocidade será indicada consoante o tipo de compasso. Assim, um compasso de  $\frac{3}{4}$  será definido à semínima enquanto o compasso de  $\frac{6}{8}$  será definido à colcheia e assim sucessivamente, sempre em bpm (batimentos por

minuto). Há ainda uma situação que poderá complementar esta solução: quando for processado o sinal adquirido, este será mostrado ao utilizador, podendo ele alterar o tempo através do alargamento ou contracção dos compassos na partitura. Será depois necessário recalcular as figuras de cada nota para aquelas medidas de compasso, algo que será feito automaticamente pelo software. Haverá também possibilidade de alterar a velocidade por um múltiplo de 2 automaticamente, tornando a modificação mais simples. Esta segunda opção será especialmente indicada para quando a primeira solução é utilizada e a transcrição foi mal efectuada, pois este será um dos erros mais comuns e de fácil solução. A confusão do tempo por múltiplos de 2 é vulgarmente constatada inclusivamente entre humanos daí que a sua solução em programas computadorizados passe também pela decisão humana.

Foi estudada a utilização de um reconhecimento automático do tempo da gravação, uma técnica ensaiada já desde meados dos anos 90 pelos holandeses Peter Desain e Henjkan Honing (8) mas concluiu-se que a sua adaptação a este programa em concreto poderia ser uma mais-valia numa fase posterior, uma vez que o resultado final pode igualmente ser obtido de forma mais simples com qualquer uma das soluções anteriores - ou ambas. Este reconhecimento seria feito pela intensidade da reprodução, reconhecendo o tempo forte do compasso como o primeiro tempo. Os tempos restantes seriam os tempos fracos que serviriam para determinar o tamanho do compasso. (5)

Deverá ser possível a eliminação de partes do sinal captado antes da transcrição de forma a otimizar o tempo de processamento que é proporcional ao tamanho do sinal a transcrever.

A determinação da duração de cada nota terá de ser determinada como pré-processamento da transcrição, para se saber que figura musical aplicar em cada nota a transcrever. Para isso é necessário fazer uma amostragem de do tempo base em caso de compasso binário e no caso de compasso ternário. Será com base nessa duração que será escolhida a figura musical para cada nota. Se a nota mudar por um intervalo inferior a metade da amostragem e voltar novamente à nota inicial o programa deve ignorar essa alteração pois deve tratar-se de um erro de aquisição. Deve, no entanto, mostrar um aviso aquando da transcrição.

O ponto seguinte é dispensável numa primeira versão do programa, podendo ser considerado um preciosismo mas será indicado para uma possível segunda fase do trabalho, tal como está definido adiante nas metas e metodologias de trabalho. Ao fazer um salto na melodia, a transição é feita com um portamento, i.e., a nota é precedida por outra(s) imediatamente acima ou abaixo, sequencialmente. Isto pode ser utilizado como forma de expressão caso no qual será muito perceptível e com um grande alcance de notas, ou erro de interpretação em que deverá ser eliminado. Poderá ainda tratar-se de uma ornamentação (trilo, trema, apoggiatura, acciatura, mordentes...). Consoante o tempo pré-determinado no metrónomo serão apenas consideradas as notas cujo tempo seja o mínimo possível. A amostragem deverá ser  $\frac{1}{2}$  ou  $\frac{1}{4}$  do tempo da menor figura possível para o tempo em questão. Por exemplo, se no metrónomo são indicados 60bpm para a semínima, a figura menor (semi-fusa) vale  $\frac{1}{16} \times 60bpm = 7,5bpm = 0,125seg$ . O tempo da amostragem será 0,0625 segundos ou 0,0313 segundos. A escolha entre estes valores dependerá de alguma experimentação.

## **2.2 Contextualização e Objectivos**

A SEEGNAL Research é uma empresa criada em 2004 como resultado de um spin-off da Faculdade de Engenharia da Universidade do Porto e do INESC-Porto. Um dos frutos desta empresa foi o “Singing Studio”, um programa de reconhecimento melódico que permite gravar e reconhecer sons musicais que podem posteriormente ser transportados para MIDI. Este software servirá de base para o trabalho a efectuar, sendo necessário aplicar-lhe a funcionalidade de conversão das notas reconhecidas para uma partitura, acompanhada de uma divisão temporal por cada nota que será necessário implementar. Após a implementação da partitura e conversão das notas incluir-se-á um processo de decisão de clave e armação de clave assim como do tipo de compassos seguindo-se a possibilidade de alterações mínimas à partitura protagonizando alguma correcção que seja urgente fazer. Terá de se incluir também um controlo de tempo (metrónomo) e, possivelmente, de precisão da amostragem.



## Capítulo 3

# Ferramentas e Metodologias a adoptar

### 3.1 Ferramentas de trabalho

O Singing Studio da SEEGNAL será o software base da dissertação. Este software utiliza o framework Qt para a criação dos efeitos gráficos e a sua compilação foi feita utilizando a versão 2005 do Microsoft Visual Studio em Windows. Utilizar-se-ão a princípio estas mesmas ferramentas apesar de se prever a experimentação de outras mais avançadas através da migração do programa para outras versões mais recentes destas ferramentas. Tal só não se toma ainda como certo devido ao enorme tempo e conhecimento necessário para efectuar a dita migração. A programação é feita utilizando a linguagem C e algumas bibliotecas específicas do programa. O grafismo é trabalhado no Qt Designer utilizando o Qt 4.7. Foi já feita uma análise do Qt com compilações básicas e alguma literatura, de forma a compreender o seu modo de funcionamento. O Qwt é uma biblioteca utilizável no Qt que contém componentes GUI e classes de utilização que são úteis principalmente para programas de índole técnica. Contém por exemplo uma ferramenta de gráfico 2D, escalas, termómetros, rodas, botões deslizantes e outro tipo de controlos que permitem mostrar valores dos mais diversos tipos. Será frequentemente utilizada nas modificações necessárias.

### 3.2 Metas e metodologia de trabalho

A meta principal desta dissertação será adicionar ao Singing Studio a funcionalidade de transcrição das notas obtidas para uma pauta musical, implementando um algoritmo que determine a sua duração e que consiga adaptá-la à partitura final com as especificações de compasso que o utilizador especificar. Terá de incluir as funcionalidades extra supra-indicadas: metrónomo e de escolha de compasso. O Qt Designer será aqui utilizado na criação do ambiente gráfico associado ao processo como a pauta, notas, pausas e demais símbolos musicais assim como botões deslizantes de velocidade de metrónomo, selecção de tipo de compasso e outros que sejam necessários. Numa segunda fase serão criados mecanismos de melhoria da transcrição como uma variação da

sensibilidade indicada pelo utilizador através de um botão deslizante. A ideia é limitar o erro obtido através da aproximação estatística da nota captada. Ao gravar um sinal audio, o software irá aproximar a frequência fundamental da nota que recebeu a uma de frequência normalizada. Se a frequência antes da aproximação ficar no limiar da decisão entre 2 notas ou se for afectada por um vibrato o programa de reconhecimento identificará uma alteração oscilante e extremamente rápida que dificilmente será uma reprodução fiável da realidade. A barra deslizante da sensibilidade permitiria seleccionar o intervalo de amostragem a considerar que por sua vez a duração mínima de uma nota. Se as notas erroneamente criadas tiverem uma duração inferior a essa devem ser ignoradas na transcrição, evitando este erro. Pela mesma razão, evitam-se assim também pequenos erros resultantes de oscilações involuntárias no traçado melódico, ou seja, pequenas desafinações que levam ao reconhecimento de notas diferentes quando na verdade se trata da mesma.

A fase final leva-nos à edição posterior da partitura transcrita. Se estas alterações forem de pequena dimensão deve ser possível fazê-las no próprio programa. Por alterações de pequena dimensão entende-se a correcção de alguma nota melódica ou ritmicamente ou a alteração do tipo de compasso. Se, no entanto, essas alterações forem mais significativas, devem ser efectuadas num programa apropriado como Encore, Finale ou Sibelius. Para isso é necessário que o software em utilização consiga exportar o seu resultado num tipo de ficheiro indicado com uma formatação própria, diferente em cada um dos objectivos. Deixa-se ainda uma proposta futura que viabiliza a optimização da afinação. Ainda que os intervalos da entoação sejam afinados, se o cantor não tiver uma primeira nota com uma correcta afinação, todo o reconhecimento poderá ser errado, o que pode ser corrigido com um simples pitch shift comandado por um botão deslizante. Outra ideia seria realizar uma análise das notas entoadas de forma a descobrir a tonalidade que deve figurar na armação de clave, e fazê-lo de forma automática. Teria de ser feita uma análise às notas que se repetiam mais vezes e ver que tonalidade contém essas mesmas notas obtendo, a principio, uma boa aproximação.

Segue-se, na figura 3.1 um diagrama de Gantt que esboça uma possível calendarização o trabalho durante o semestre em questão. A azul claro está o trabalho geral e a azul escuro o da primeira fase, a essencial. Em caso de adiantamento, passa-se à fase seguinte. A fase 2 está a verde e a fase 3 em laranja, todas devidamente explicitadas na figura.

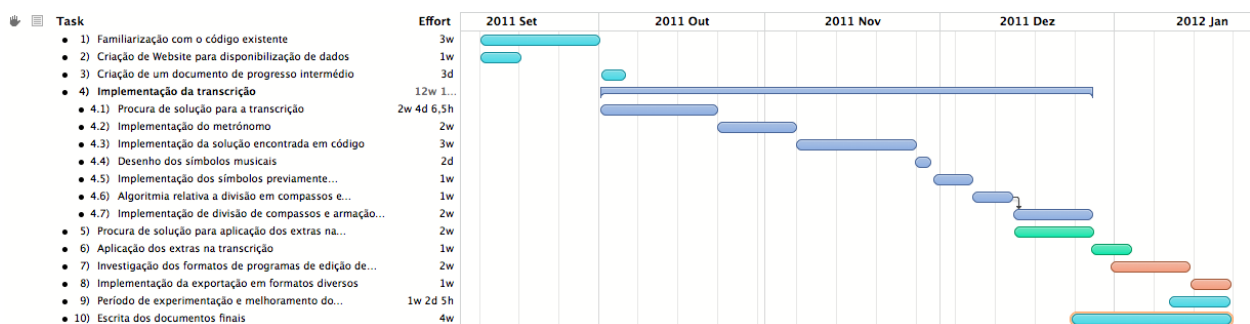


Figura 3.1: Diagrama de Gantt com calendarização dos trabalhos

## Capítulo 4

### Resumo

A evolução da tecnologia permitiu o acesso cada vez mais facilitado à música mas, no entanto, não levou a inovações revolucionárias no que toca à escrita musical. Analisando o software existente conclui-se que não há um programa de computador que facilite a transcrição musical que o faça satisfazendo as exigências de um músico, apenas servindo para fins de experimentação e de recreio. Partindo do Singing Studio, um software de identificação de notas musicais a partir da gravação de um ficheiro de som, pretende-se juntar-lhe uma funcionalidade de transcrição para pauta musical de forma automática que consiga providenciar alguma fiabilidade no processo, podendo, numa fase final, acrescentar-lhe outras funcionalidades de exportação para programas profissionais de edição de partituras amplamente utilizados. Tal software culminaria numa utilização evidentemente prática das técnicas de processamento de sinais e teria a máxima utilidade aplicada ao ensino da música e inclusivamente na utilização profissional de criadores de música, acelerando o processo de escrita e abrindo caminho à criatividade.





# Referências

- [1] Luís L Henrique. *Acústica musical*. Fundação Calouste Gulbenkian, Lisboa, 2002. Luís L. Henrique 26 cm inclui 1 CD-ROM.
- [2] Daniel J. Levitin. *This is your brain on music : the science of a human obsession*. Dutton, New York, N.Y., 2006. 2006009055 Daniel J. Levitin. ill. ; 24 cm. Includes bibliographical references (p. [271]-300) and index. Introduction: I love music and I love science– why would I want to mix the two? – What is music? : from pitch to timbre – Foot tapping : discerning rhythm, loudness, and harmony – Behind the curtain : music and the mind machine – Anticipation : what we expect from Liszt (and Ludacris) – You know my name, look up the number : how we categorize music – After dessert, Crick was still four seats away from me : music, emotion, and the reptilian brain – What makes a musician? : expertise dissected – My favorite things : why do we like the music we like? – The music instinct : evolution’s 1 hit.
- [3] D. M. Huber. *The MIDI manual: a practical guide to MIDI in the project studio*. Focal Press, 2007. 2007280783, <http://books.google.com/books?id=GfHZwBwZuKIC>.
- [4] E. C. Axford. *Song sheets to software: a guide to print music, software, and web sites for musicians*. Scarecrow Press, 2004. 2004002488, <http://books.google.com/books?id=mNv9kAuTDTIC>, G - Reference, Information and Interdisciplinary Subjects Series.
- [5] H. A. Dourado. *Dicionário de termos e expressões da música*. Editora 34, 2004. 2006337075.
- [6] M. Mascarenhas, B. Cardoso. *Curso Completo de Teoria Musical e Solfejo*. Irmãos Vitale, 1996. 9788574070988, volume 2.
- [7] J. G. Sacramento. *Notação Musical*. José Sacramento, 2007. 9789899528802.
- [8] Peter Desain, Henkjan Honing. *Music, mind, and machine : studies in computer music, music cognition, and artificial intelligence*. Thesis, Amsterdam, 1992. Kennistechologie. 92231276, 330 p., Peter Desain and Henkjan Honing. ill. ; 24 cm. Includes bibliographical references.
- [9] P. Martins Silva. *Elementos de acústica musical*. Laboratório Nacional de Engenharia Civil, Lisboa, 1989. 231 p., P. Martins da Silva il. 30 cm.