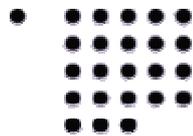


# Neuronale Netze und Musik

von  
M. Serhat Cinar  
AI 11030409

Ein Referat  
über die Doktorarbeit  
„Lernen musikalischer Strukturen und  
Stile mit neuronalen Netzen“  
von Dominik Hörnel

für das Wahlpflichtfach  
Neuronale Netze II  
bei Prof. Dr. H. Westenberger  
Wintersemester 2003-2004



Fachhochschule Köln  
Campus Gummersbach

## Inhalt

Abbildungsverzeichnis.....	3
1 Einleitung .....	4
1.1 Das Thema.....	4
2 Musikalische Grundlagen – Die Elemente der Musik.....	4
2.1 Der Ton.....	4
2.2 Die Oktave.....	5
2.3 Das Tonsystem.....	5
2.4 Der Notenschlüssel in der Notation.....	6
2.5 Der Rhythmus.....	6
2.6 Die Harmonien.....	6
2.7 Die Akkorde.....	8
2.8 Aufbau eines Chorals.....	8
3 Übergeordnete melodische Strukturen .....	9
3.1 Das Motiv.....	9
3.2 Die Hauptdreiklänge.....	10
4 Das Grundproblem – Sequenzen und Neuronale Netze.....	10
4.1 STM Modelle .....	10
4.1.1 Feedforward-Netze mit fester Kontextlänge.....	10
4.1.2 Rückgekoppelte Netze – Jordan-Netze.....	11
4.1.3 Rückgekoppelte Netze – Elman-Netze.....	11
4.1.4 Rückgekoppelte Netze – CONCERT-Netz.....	11
4.1.5 Rückkopplung durch Gedächtnisneuronen.....	12
4.1.6 Mehrschichtige STM Modelle.....	12
5 Die Aufgaben.....	12
5.1 Komposition von Melodieumspielungen – Melonet und Jazznet.....	12
5.1.1 Melonet .....	13
5.1.2 Jazznet.....	18
5.2 Melodiekomposition – Melogenet.....	20
5.2.1 Melogenet .....	22
5.3 Choralharmonisierung – Harmonet und Harmogenet.....	25
5.3.1 Harmonet.....	25
5.3.2 Harmogenet.....	27
5.4 Stilanalyse und Stilerkennung.....	28
6 Literaturverzeichnis.....	30

## Abbildungsverzeichnis

Abbildung 1: Ein Sinuston.....	4
Abbildung 2: Eigenschaften verschiedener Töne.....	5
Abbildung 3: Saitenverhältnisse .....	5
Abbildung 4: Tastatur eines Keyboards.....	6
Abbildung 5: Violin- und Bassschlüssel.....	6
Abbildung 6: Übersicht der Notendauern in der Notation.....	6
Abbildung 7: C-Dur Tonleiter.....	7
Abbildung 8: Bezeichnungen der verschiedenen Intervalle .....	7
Abbildung 9: D-Dur Tonleiter .....	7
Abbildung 10: Choralvariation von J. Pachelbel „Christus, der ist mein Leben“ .....	8
Abbildung 11: Die ersten beiden Takte der Allegro con brio aus der 5. Symphonie Beethovens...9	9
Abbildung 12: Darstellung der Tonfolge als Gestaltabbildung.....	9
Abbildung 13: Paradigmatische Gruppe des Beethoven Motivs.....	9
Abbildung 14: Die Hauptdreiklänge der c-Dur Tonleiter.....	10
Abbildung 15: Das Modell eines Jordan-Netzes .....	11
Abbildung 16: CONCERT-Netz mit rückgekoppelter verborgener Schicht.....	11
Abbildung 17: Schematische Darstellung eines zweistufigen STM Modells .....	12
Abbildung 18: Struktur von Melonet .....	14
Abbildung 19: Kohonen-Netz aller Motive aus einer Choralumspielung .....	15
Abbildung 20: Komplementäre Intervallkodierung .....	16
Abbildung 21: Ergebnisse der verschiedenen Netze zum Präludium V von J.S. Bach .....	17
Abbildung 22: Komitee von mehreren Netzen .....	17
Abbildung 23: Ausschnitt einer Melodievariation über die Chormelodie „Werde munter, mein Gemüte“ vom Melonet-Komitee.....	18
Abbildung 24: Hierarchische Bestimmung einer Teilphrase durch Jazznet .....	19
Abbildung 25: Aufsummierte Richtungswerte für Jazzrule, Jazznet und Parker.....	19
Abbildung 26: Verteilung der Inside, Outside und Leading Achteltöne.....	20
Abbildung 27: Beispielimprovisationen von Jazzrule, Jazznet und Charlie Parker .....	20
Abbildung 28: „Analysis by synthesis“ in Melogenet .....	21
Abbildung 29: Analyse und funktionale Abhängigkeiten in Melogenet.....	22
Abbildung 30: Stilerkennung eines chinesischen Volksliedes durch automatische Fitnessbewertung .....	23
Abbildung 31: Melodien, die für das Hörexperiment verwendet wurden und deren Fitness.....	24
Abbildung 32: Ergebnis des Experiments, aufgeteilt nach Teilnehmer (o=Teilnehmer tippt auf Original).....	24
Abbildung 33: Aufbau von Harmonet.....	25
Abbildung 34: Harmonische Tonkodierung für c-Dur .....	26
Abbildung 35: Aufbau von Harmogenet .....	27
Abbildung 36: Stilerkennungsprozess .....	28
Abbildung 37: Stilerkennungsgüte für drei Netze (Bach-, Reger- und Scheidt-Stil) .....	28
Abbildung 38: Harmonische Erwartungen des Bach- und Scheidt-Netzes für den Choral .....	29

# 1 Einleitung

Dieses Referat beschäftigt sich hauptsächlich mit der von Dominik Hörnel verfassten Doktorarbeit mit dem Titel „Lernen musikalischer Strukturen und Stile mit neuronalen Netzen“. Bis auf einige eigene Ausführungen, vor allem im einleitenden Kapitel über die musikalischen Grundlagen, stammt der Inhalt aus der genannten Doktorarbeit. Daher kann dieses Referat als eine „Zusammenfassung“ des Buches interpretiert werden.

## 1.1 Das Thema

„Ziel der Arbeit ist die Modellierung musikalischer Strukturen und das Aufspüren stiltypischer Merkmale aus Musikbeispielen mit Hilfe neuronaler Netze, sowie deren Anwendung auf praxisrelevante musikalische Problemstellungen:

das Erfinden melodischer Umspielungen nach Art barocker Choralvariationen [...] und Jazzimprovisationen [...], die Vervollständigung von Melodiefragmenten und die Harmonisierung von Chormelodien [...].“<sup>1</sup>

Die Arbeit befasst sich somit mit der Analyse und Erkennung von musikalischen Stilen in Musikstücken.

Mit Stilerkennung ist hier gemeint, dass anhand gelernter Musikstile aus Stücken, für ein neues Stück, dessen Komponist unbekannt ist, der Komponist bestimmt und damit einem Stil zugeordnet wird.

## 2 Musikalische Grundlagen – Die Elemente der Musik

Für Nichtmusiker (wie mich) soll im Voraus einiges an Grundwissen zur Musiklehre vermittelt werden, damit die im weiteren Verlauf angewendete Analyse von Kompositionen verständlicher und auch besser bewertbar wird.

### 2.1 Der Ton<sup>2</sup>

Als Ton werden regelmäßige, sinusförmige Schwingungen bezeichnet. Sind die Schwingungen nicht sinusförmig, spricht man vom Klang. Liegt keine Regelmäßigkeit oder Periodizität vor, so spricht man von einem Geräusch. Im folgenden beschäftigen wir uns hauptsächlich mit Tönen.

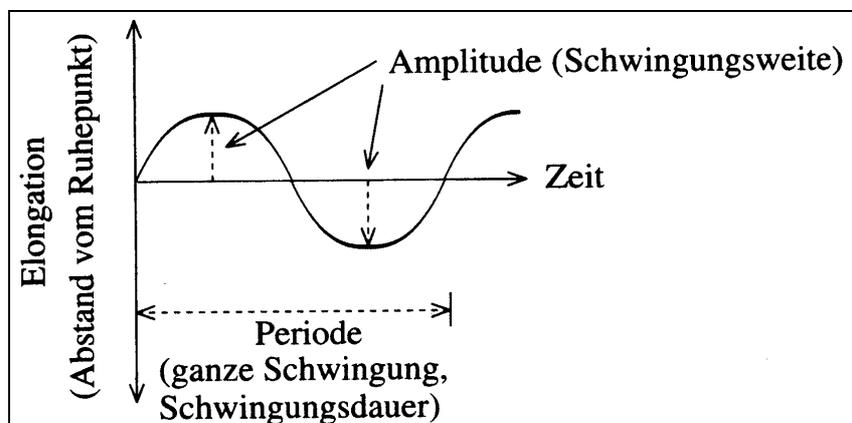


Abbildung 1: Ein Sinuston<sup>3</sup>

<sup>1</sup> [Hörnel 00] aus der Zusammenfassung

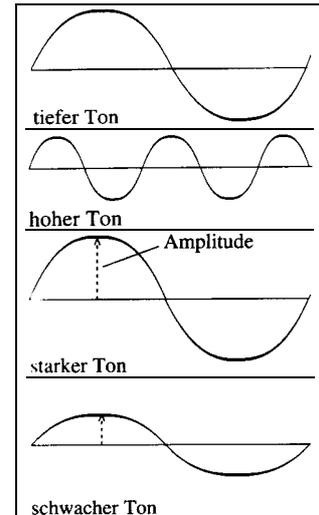
<sup>2</sup> Zusammengefasst aus [Ziegenrucker 97]

Geräusche kommen in klassischer Musik hauptsächlich als Perkussion vor.

Zu den wichtigsten Eigenschaften eines Tones zählen Höhe, Stärke und Farbe.

Die Tonhöhe wird durch die Anzahl der Schwingungen pro Sekunde bestimmt (viele Schwingungen pro Sekunde ergibt einen hohen Ton), die Stärke durch die Höhe der Amplitude (eine hohe Amplitude ergibt einen starken Ton).

Die Farbe eines Tones wird vor allem durch die Obertöne bestimmt. Jedes Instrument besitzt auch bei gleichem Ton eine eigene (unverkennbare und typische) Klangfarbe. Diese entsteht dadurch, dass die Instrumente keine reinen Sinustöne erzeugen, sondern Sinustöne mit feinen weiteren überlagernden Schwingungen, den Obertönen.



**Abbildung 2:**  
**Eigenschaften**  
**verschiedener Töne<sup>4</sup>**

## 2.2 Die Oktave

Zupft man eine (gestimmte) Saite, so ertönt ein Grundton mit einer bestimmten Frequenz, z.B. c. Halbiert man die Saite, verdoppelt sich die Frequenz. Beide Töne klingen dann gleich, mit dem Unterschied, dass der Ton mit der höheren Frequenz auch höher klingt, z.B. hohes c. Der Abstand zwischen zwei Tönen mit doppelter Frequenz wird als Oktave bezeichnet.

schwingende Saite	$\frac{1}{1}$ = Grundton (C)
	$\frac{1}{2}$ = Oktave (c)
	$\frac{1}{3}$ = Oktave + Quinte (g)
	$\frac{1}{4}$ = Doppeloktave (c')
usw.	
<b>abgeleitete Intervallverhältnisse:</b>	
reine Oktave	1 : 2
reine Quinte	2 : 3
reine Quarte	3 : 4
große Terz	4 : 5
kleine Terz	5 : 6
große Sekunde	8 : 9
kleine Sekunde	15 : 16

**Abbildung 3: Saitenverhältnisse<sup>5</sup>**

## 2.3 Das Tonssystem

Als Tonssystem bezeichnet man die Unterteilung einer Oktave in einzelne (Halb-) Noten. In dem uns bekannten Tonssystem, das typisch für Europa ist, wird eine Oktave in 12 Halbtöne unterteilt. Diese Unterteilung basiert auf Pythagoras<sup>6</sup>, der die 12 Halbtöne anhand der Saitenverhältnisse berechnete (z.B. Oktave = 1:2, Sekunde = 1 ganzer Ton = 15:16) (siehe Abbildung 3: Saitenverhältnisse).

<sup>3</sup> Aus [Ziegenrucker 97] S.10.

<sup>4</sup> Aus [Ziegenrucker 97] S.10/11.

<sup>5</sup> Aus [Ziegenrucker 97] S.12.

<sup>6</sup> Pythagoras, griechischer Philosoph und Mathematiker, 582-496 v. Chr.

Hieraus ergibt sich die Unterteilung einer Oktave in 7 ganze und 5 halbe Töne (insgesamt 12 Halbtöne). Entsprechend ist auch die Tastatur eines Klaviers oder eines Keyboards aufgeteilt<sup>7</sup>. Die weißen Tasten sind die Grundtöne, also die ganzen Noten, die schwarzen Tasten sind die Halbtöne.

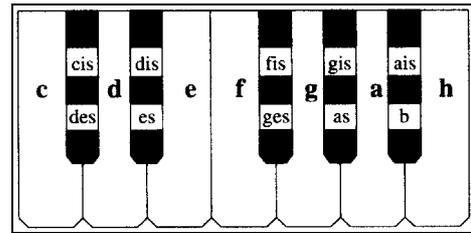


Abbildung 4: Tastatur eines Keyboards<sup>8</sup>

## 2.4 Der Notenschlüssel in der Notation

In der Notation von Noten wird ein Notenschlüssel benutzt, um die Noten einfacher zu erkennen. Hierbei gibt es den Violinschlüssel, der auch g-Schlüssel genannt wird, da die Schleife des Schlüssels die Linie für die Note g umschließt, und für tiefere Töne den Bassschlüssel, der auch f-Schlüssel genannt wird, da die Doppelpunkte die Linie für die Note f markieren.

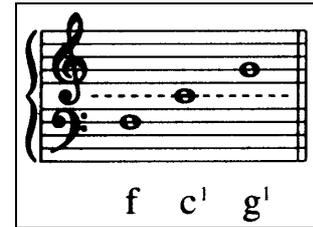


Abbildung 5: Violin- und Bassschlüssel<sup>9</sup>

## 2.5 Der Rhythmus

Der zeitliche Ablauf der Musik wird u.a. durch die Gliederung der Notendauern bestimmt. Dabei wird vom Takt, einer Zeiteinheit, ausgegangen. Jede Note kann über einen ganzen, einen halben, einen viertel oder einen achtel Takt gespielt werden. Hierbei sind auch Pausen mit denselben Längen möglich. In der Notation erkennt man die Notendauer am ausgefüllten Notenknopf und an dem Schweif der Note.

Name	Notenbild	Pausenbild	Name
Ganze Note	o	—	Ganze Pause
Halbe Note	♪ ♩	—	Halbe Pause
Viertelnote	♪ ♪	—	Viertelpause
Achtelnote	♪ ♫	—	Achtelpause
Sechzehntelnote	♪ ♯	—	Sechzehntelpause
Zweiunddreißigstelnote	♪ ⁄	—	Zweiunddreißigstelpause
Vierundsechzigstelnote	♪ ⁄	—	Vierundsechzigstelpause

Abbildung 6: Übersicht der Notendauern in der Notation<sup>10</sup>

## 2.6 Die Harmonien

„Der Begriff Harmonie umfasst jedes räumliche miteinander von Tönen, die Ordnung der Zusammenklänge. Gegenstand der Harmonielehre ist neben dem Aufbau der Akkorde insbesondere die Verbindung der Klänge zu musikalisch logischen Folgen.“<sup>11</sup>

<sup>7</sup> In einigen Ländern werden für die ganzen Töne die Bezeichnungen Do, Re, Mi, Fa, So, La, Si, Do benutzt.

<sup>8</sup> Aus [Ziegenrucker 97] Seite 34.

<sup>9</sup> Aus [Ziegenrucker 97] Seite 25.

<sup>10</sup> Aus [Ziegenrucker 97] Seite 40.

<sup>11</sup> [Ziegenrucker 97] Seite 131.

Zur Bildung von Harmonien benötigt man die *Grundtonreihe* als Grundlage. Nimmt man c als Grundton, so ergibt sich die c-Dur Grundtonreihe c, d, e, f, g, a, h.



Abbildung 7: C-Dur Tonleiter<sup>12</sup>

Dabei gilt es die auf der Tastatur erkennbaren Abstände bezüglich der Halbtöne einzuhalten. Beispielsweise liegt zwischen dem ersten und dem zweiten Ton der c-Dur Leiter ein Halbton, das cis. Zwischen dem zweiten Ton (d) und dem dritten Ton (e) existiert wieder ein Halbton (dis). Zwischen dem dritten Ton (e) und dem vierten Ton (f) existiert hingegen kein Halbton.

Den Abstand zwischen zwei Tönen einer Tonreihe nennt man Sekunde (Sekunde: von der Zahl 2 abgeleitet, da mit dem ersten Ton zusammen der Abstand 2 Töne beträgt: c - d). Es gibt auch Bezeichnungen für den Abstand von c zu e, die Terz (Terz: von der Zahl drei c - d - e). Je nachdem, ob noch ein Halbton dazwischen liegt oder nicht, unterscheidet man die kleine Sekunde von der großen Sekunde.

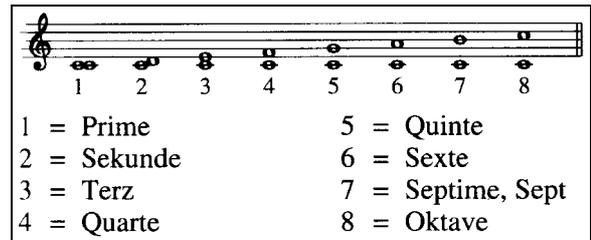


Abbildung 8: Bezeichnungen der verschiedenen Intervalle<sup>13</sup>

Erster Ton - zweiter Ton	c - d	Halbton dazwischen cis	Grosse Sekunde
Zweiter Ton - dritter Ton	d - e	Halbton dazwischen dis	Grosse Sekunde
Dritter Ton - vierter Ton	e - f	Kein Halbton dazwischen	Kleine Sekunde

Tabelle 1: C-Dur Tonleiter und die Intervallbezeichnungen

Diese Abstände sind zu berücksichtigen, wenn man die d-Dur Tonleiter bilden möchte. Die d-Dur Tonleiter beginnt mit dem Grundton d. Der zweite Ton der d-Dur Leiter ist der Ton e. Zwischen erstem und zweitem Ton muss wieder ein Halbton (dis) liegen. Zwischen zweitem und drittem Ton muss hier auch wieder ein halber Ton dazwischen liegen. Da jedoch zwischen e und f kein halber Ton liegt, muss f als Halbtontschritt benutzt werden, was dazu führt, dass der dritte Ton der d-Dur Tonleiter statt dem f das fis ist.



Abbildung 9: D-Dur Tonleiter<sup>14</sup>

Erster Ton - zweiter Ton	d - e	Halbton dazwischen dis	Grosse Sekunde
Zweiter Ton - dritter Ton	e - fis	Halbton dazwischen f	Grosse Sekunde
Dritter Ton - vierter Ton	fis - g	Kein Halbton dazwischen	Kleine Sekunde

Tabelle 2: Die D-Dur Tonleiter und die Intervallbezeichnungen

Anhand der Intervalle kann man also zu jedem Grundton die Dur Tonleiter bilden. Gleichermassen kann man auch andere Tonleitern bezüglich anderer harmonischer Mengen bilden. Die Molltonleiter orientiert sich bei der Bildung der Tonleiter an den Intervallen der Tonleiter a-Moll: a, h, c, d, e, f, g. Hier liegt z.B. zwischen dem zweiten Ton und dem dritten Ton der Tonleiter kein Halbton, also handelt es sich hierbei um eine kleine Sekunde. Entsprechend lautet der Anfang der c-Moll Leiter c, d, dis, da zwischen d und dem nächsten Ton der Leiter eine kleine Sekunde (kein Halbton) liegt.

<sup>12</sup> Aus [Ziegenrucker 97] Seite 104.

<sup>13</sup> Aus [Ziegenrucker 97] Seite 77.

<sup>14</sup> Aus [Ziegenrucker 97] Seite 104.

## 2.7 Die Akkorde

Unter Akkorden versteht man das gleichzeitige Spielen mehrerer Noten, die einen Gesamtklang erzeugen. Am häufigsten tritt der Dreiklang auf, das zusammen spielen von drei verschiedenen Tönen. Da nicht jede beliebige Kombination von drei Tönen harmonisiert (gut klingt), gibt es ein Regelwerk, die Harmonielehre, die sich mit dem Aufbau von harmonischen Akkorden beschäftigt. Spielt man z.B. die Töne c, cis, d, dis, e gleichzeitig, entsteht ein unheimlicher, unharmonischer Klang, der höchstens in Horrorszene von Filmen Verwendung findet<sup>15</sup>, aber sehr selten in der klassischen Komposition.

Akkorde werden erzeugt, indem der erste, der dritte und der fünfte Ton einer Tonleiter gleichzeitig gespielt werden.

In der c-Dur Tonleiter sind dies die Töne c, e und g. Gemeinsam ergeben sie den c-Dur Akkord.

Weitere Akkordarten sind Vierklänge, wobei meist ein Dreiklang durch eine weitere, tiefe Note der entsprechenden Tonleiter begleitet wird.

## 2.8 Aufbau eines Chorals

Ein klassischer Choral (Chorgesang, meist in der Kirche gesungen, wie z.B. die gregorianischen Kantenchöre) baut auf einer Melodie aus einzelnen Tönen auf, die auf einer Tonleiter liegen. Zu jedem Ton wird eine Harmonie bestimmt, aus der dann Akkorde abgeleitet werden. Die Akkorde begleiten klanglich die Melodie.

Schließlich gibt es noch die Melodieumspielung, welche aus einzelnen Tönen besteht, die in der „Nähe“ der Melodie gespielt werden. Während die Töne der Melodie innerhalb eines Taktes nur ein oder zweimal wechseln, wechseln die Töne der Melodieumspielung mehrfach während eines Taktes. So ergibt sich insgesamt ein voller und harmonischer Klang.

Melodie

Variation  
(Umspielung & Akkorde)

Harmonik

G: T T D3 T S3 D T S T3 D5 T D T

Abbildung 10: Choralvariation von J. Pachelbel<sup>16</sup> „Christus, der ist mein Leben“<sup>17</sup>

<sup>15</sup> Dieser unharmonische Mehrklang wurde im Film „Psycho“ von Alfred Hitchcock benutzt, wo er in einer der berühmtesten Filmszenen überhaupt, in welcher der Mörder sein erstes Opfer in der Badewanne ersticht, eine spannende, nervenaufreibende Atmosphäre eben durch seine Disharmonie erzeugt. Auch in dem Film „Der weiße Hai“ sind in einigen Szenen (Haiangriffe) Disharmonien zum gleichen Zwecke eingesetzt.

<sup>16</sup> Johann Pachelbel, dt. Komponist (1653-1706).

<sup>17</sup> Aus [Hörnel 00] S.9.

### 3 Übergeordnete melodische Strukturen

Bis jetzt wurde sehr auf die Struktur der einzelnen Töne eingegangen, wie sie z.B. miteinander harmonisieren. Es gibt im Aufbau eines Liedes auch übergeordnete Strukturen. Dies bezieht sich vor allem auf den zeitlichen Ablauf der Töne und Akkorde.

#### 3.1 Das Motiv

Um das Motiv anschaulich darzustellen wird im Folgenden der erste Satz, Allegro con brio, der 5. Symphonie von Beethoven<sup>18</sup> betrachtet. In den Tönen aus der 5. Symphonie taucht immer wieder die Tonfolge g ,g , g , d auf, die dem Stück seine unverkennbare „Note“ verleiht.

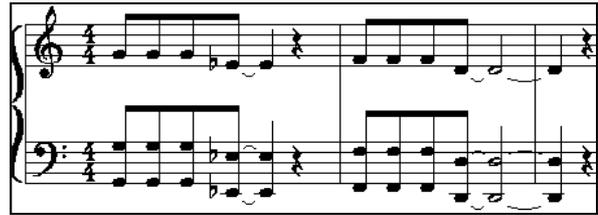


Abbildung 11: Die ersten beiden Takte der Allegro con brio aus der 5. Symphonie Beethovens

Entscheidend bei dieser Tonfolge sind nicht die einzelnen Noten (g oder d), sondern die absteigende Struktur der Noten. Dieses relative Verhältnis aufeinander folgender Töne wird als Motiv bezeichnet. Erkennbar wird die

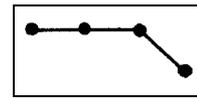


Abbildung 12: Darstellung der Tonfolge als Gestaltabbildung<sup>19</sup>

Unabhängigkeit des Motivs von konkreten Tönen, wenn man die ersten beiden Takte der 5. Symphonie betrachtet. Das Motiv wird im zweiten Takt erneut gleich gespielt, jedoch hier mit anderen Noten.

Motive werden zu so genannten paradigmatischen Gruppen zusammengefasst. Dabei wird ein Motiv und seine Spiegelungen an den Achsen zu einer Gruppe zusammengefasst.

Betrachtet man die 5. Symphonie Beethovens, so fällt auf, dass das gesamte Stück aus fast nur einer Motivgruppe besteht. Dies gilt auch als das besondere dieser Symphonie (weswegen sie hier auch als anschauliches Beispiel sehr gut geeignet ist). Das Motiv taucht hier immer wieder auf. (Anscheinend wollte Beethoven hier demonstrieren, was man alles aus nur einem Motiv machen kann). Dies ist aber nicht immer so. In anderen Stücken gibt es auch mehrere, verschiedene Motive.

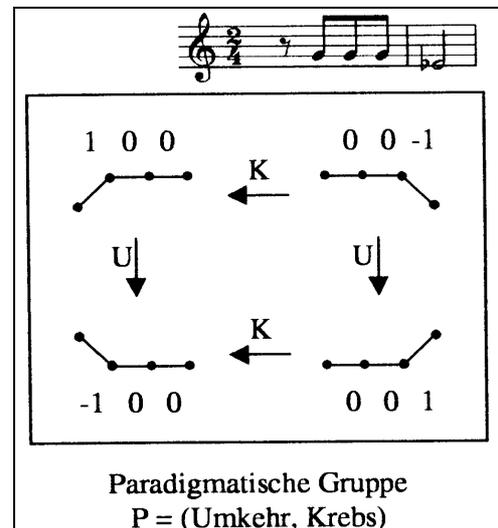


Abbildung 13: Paradigmatische Gruppe des Beethoven Motivs<sup>20</sup>

<sup>18</sup> Ludwig van Beethoven, dt. Komponist (1770-1827).

<sup>19</sup> Aus [Hörnel 00] S.72.

<sup>20</sup> Aus [Hörnel 00] S.72.

### 3.2 Die Hauptdreiklänge

In der c-Dur Tonleiter ist der Akkord c-e-g nicht der einzige, der benutzt werden kann. Auch f-a-c und g-h-d sind harmonische Dreiklänge. Sie werden entsprechend der ersten Note des Dreiklanges Akkord der ersten Stufe (c-e-g wegen c = erster Ton, auch *Tonika*), der vierten Stufe (f-a-c wegen f = vierter Ton, auch *Subdominante*) und der fünften Stufe (g-h-d wegen g = fünfter Ton, auch *Dominante*) genannt.

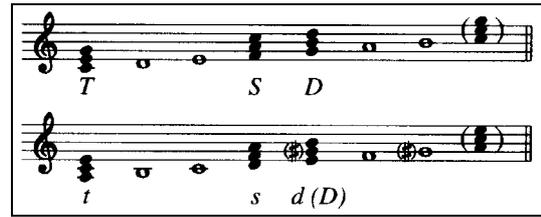


Abbildung 14: Die Hauptdreiklänge der c-Dur Tonleiter<sup>21</sup>

Eine besondere Rolle spielen diese verschiedenen Hauptdreiklänge vor allem in ihrer Abfolge.

„Die Verbindung der Hauptdreiklänge führt zur einfachen Kadenz (Grundkadenz), die durch das harmonische Gefälle, durch Spannung und Entspannung charakterisiert ist. Die Hauptdreiklänge stehen untereinander in funktionalem Verhältnis. [...] Die Tonika bildet das Zentrum des harmonischen Geschehens (tonos, griechisch = Spannung). Sie wird – bildlich betrachtet – vom Oberdominant- und Unterdominantbereich umgeben.[...] Das Spannungsgefüge der Kadenz ist durch den doppelten Quintfall charakterisiert: Die Tonika ‚fällt‘ zur Subdominante (Entfernung vom tonalen Zentrum – Spannung), die Dominante ‚fällt‘ zur Tonika (Hinführung zum tonalen Zentrum – Entspannung).“<sup>22</sup>

## 4 Das Grundproblem – Sequenzen und Neuronale Netze

Um musikalische Strukturen zu lernen, muss ein Netz mehrere dieser Aspekte berücksichtigen. Zum einen gilt es, zeitlich untergeordnete Strukturen, wie die korrekten Töne eines Akkordes zu lernen, zum anderen müssen zeitlich übergeordnete Strukturen, wie Motive oder Harmoniefolgen gelernt werden. Eine anschauliche Analogie bilden hier Texte: untergeordnete Strukturen wären hier Buchstaben, die sich zu Wörtern vereinen, während übergeordnet einzelne Wörter einen Satz ergeben.

Wie bereits im Praktikum für Neuronale Netze I, besteht die Möglichkeit, den zeitlichen Kontext der Noten durch ein Neuronales Netz mit der Fenstertechnik (Sequenzen-Lernen) zu trainieren. Abstrahiert spricht man von so genannten STM – Short Term Memory (Kurzzeitgedächtnis) Modellen.

### 4.1 STM Modelle

#### 4.1.1 Feedforward-Netze mit fester Kontextlänge

Bei Feedforward-Netzen mit fester Kontextlänge wird dem Netz mit dem zu lernenden Sequenzparameter  $s(t)$  auch der vorangegangene Kontext, z.B.  $s(t-1)$  und  $s(t-2)$ , präsentiert. Dadurch lernt das Netz, einem bestimmten Kontext  $s(t-2)$ ,  $s(t-1)$ ,  $s(t)$  einen bestimmten Wert  $s(t+1)$  zu zuordnen.

<sup>21</sup> Aus [Ziegenrücker 97] S.138.

<sup>22</sup> Aus [Ziegenrücker 97] S.138 und folgende.

### 4.1.2 Rückgekoppelte Netze – Jordan-Netze

„In Jordan-Netzen [...] setzt sich die Eingabeschicht aus zwei Teilen zusammen. der erste Teil besteht aus so genannten *Planneuronen*, die externe Eingaben erhalten und der Identifizierung einer abzurufenden Sequenz dienen. Der zweite Teil besteht aus Zustandsneuronen, zu denen die Ausgabe des Netzes zurückgeführt wird und die den aktuellen Zustand einer Sequenz repräsentiert. Durch die [...] gewichtete Verbindung der Ausgabeneuronen zu den Zustandsneuronen merkt sich das Netz einen Teil der Aktivierung des letzten Symbols.“<sup>23</sup>

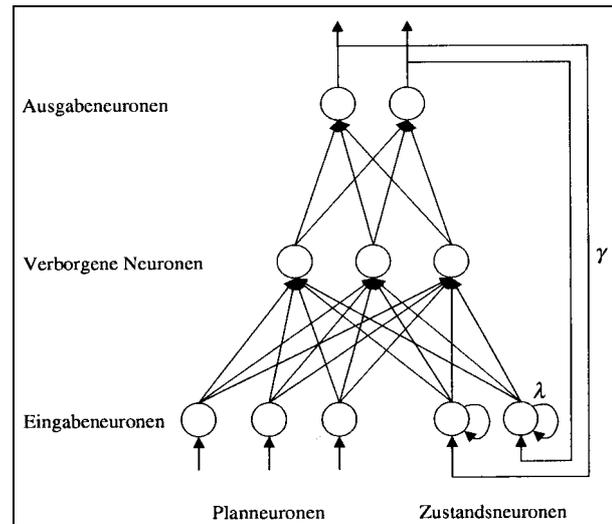


Abbildung 15: Das Modell eines Jordan-Netzes<sup>24</sup>

### 4.1.3 Rückgekoppelte Netze – Elman-Netze

„In Elman-Netzen [...] sind an Stelle der Ausgabeneuronen ein Teil der verborgenen Neuronen mit der Eingabeschicht verbunden. Diese Kontextschicht kopiert die Aktivierungen der verborgenen Schicht des vorherigen Zustands und verwendet diese zur Bestimmung des aktuellen Symbols.“<sup>25</sup>

### 4.1.4 Rückgekoppelte Netze – CONCERT-Netz

Im Kompositionssystem CONCERT von M. C. Mozer wird ein modifiziertes dreischichtiges Feedforward-Netz konstruiert, in dem die verborgenen Neuronen zu sich selbst rückgekoppelt werden.

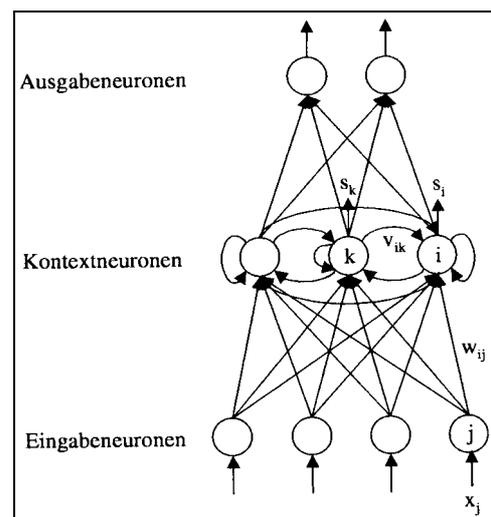


Abbildung 16: CONCERT-Netz mit rückgekoppelter verborgener Schicht<sup>26</sup>

<sup>23</sup> Aus [Hörnel 00] S.41/42.

<sup>24</sup> Aus [Hörnel 00] S.42.

<sup>25</sup> Aus [Hörnel 00] S.43.

<sup>26</sup> Aus [Hörnel 00] S.44.

#### 4.1.5 Rückkopplung durch Gedächtnisneuronen

In einem weiteren Modell wird die verborgene Schicht mit Gedächtnisneuronen erstellt. Gedächtnisneuronen verändern ihre Ausgabe erst nach jeder ganzen Sequenz, wodurch auch das Lernverfahren entsprechend angepasst werden muss. „Die Delay/Offset-Werte [mit denen bestimmt wird, wann ein Neuron eine Ausgabe produziert] hatten dabei einen großen Einfluss auf das Lernergebnis: Die Verwendung mehrerer Gedächtnisneuronen mit unterschiedlichen Delay/Offset-Werten führte meist zu besseren Ergebnissen, Gedächtnisneuronen mit gleichen Delay/Offset-Wert erwiesen sich als nicht sinnvoll.“<sup>27</sup>

#### 4.1.6 Mehrschichtige STM Modelle

Auch die einfachen STM-Netze können übergeordnete Strukturen nicht gut lernen. Daher wird von Hörnel ein mehrstufiges Modell präsentiert, bei dem ein Subnetz lokale, untergeordnete Strukturen lernt und seine Ergebnisse einem Supernetz präsentiert, der aus den Ausgaben mehrerer Subnetze eine übergeordnete Struktur lernt.

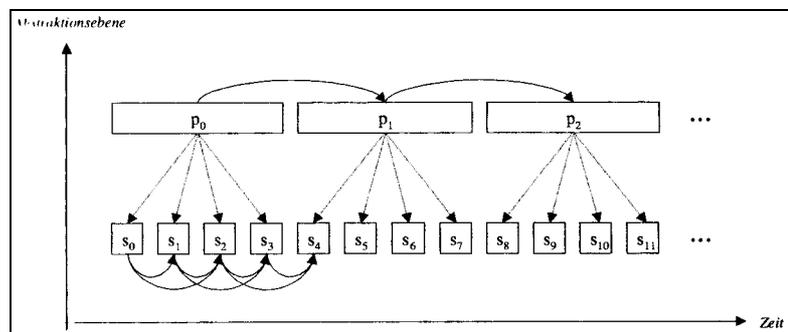


Abbildung 17: Schematische Darstellung eines zweistufigen STM Modells<sup>28</sup>

## 5 Die Aufgaben

In der Arbeit von Dominik Hörnel werden für verschiedene Aufgaben verschiedene Konzepte aufgestellt. Im Folgenden werden diese Aufgaben und ihre Lösungen dargestellt.

### 5.1 Komposition von Melodieumspielungen – Melonet und Jazznet

Wie bereits erwähnt, sind Melodieumspielungen Noten, die in der Nähe der Melodie harmonisch zu ihr die Melodie begleiten.

„Das Lernen von Melodievariationen stellt in seiner Komplexität gewissermaßen eine Vorstufe zum Lernen ganzer Melodien dar. Durch die Vorgabe einer zu umspielenden Melodie bzw. einer Harmoniefolge, über die variiert bzw. improvisiert werden soll, liegt über den auszufüllenden Zeitraum bereits eine musikalische Grundstruktur vor, an welcher sich der Komponist [...] orientieren kann. Die Schwierigkeit der Aufgabe besteht darin, eine Umspielung zu finden, die zu eben diesen Vorgaben passt, d.h. insbesondere die Abhängigkeit zwischen harmonischen und melodischen Beziehungen berücksichtigt.“<sup>29</sup>

<sup>27</sup> Aus [Hörnel 00] S.46.

<sup>28</sup> Aus [Hörnel 00] S.51.

<sup>29</sup> [Hörnel 00] S.105/106.

In der Realität muss ein Musiker in Echtzeit *improvisieren*. Daher ist dem Musiker meist die Melodie sowie der linke Kontext, also alles, was er zuvor gespielt hat, bekannt. Es gilt jeweils die nächste Note im Kontext vorauszusagen. Auf dieser Grundlage bauen auch die folgenden beiden Netze auf, die improvisieren.

### **5.1.1 Melonet**

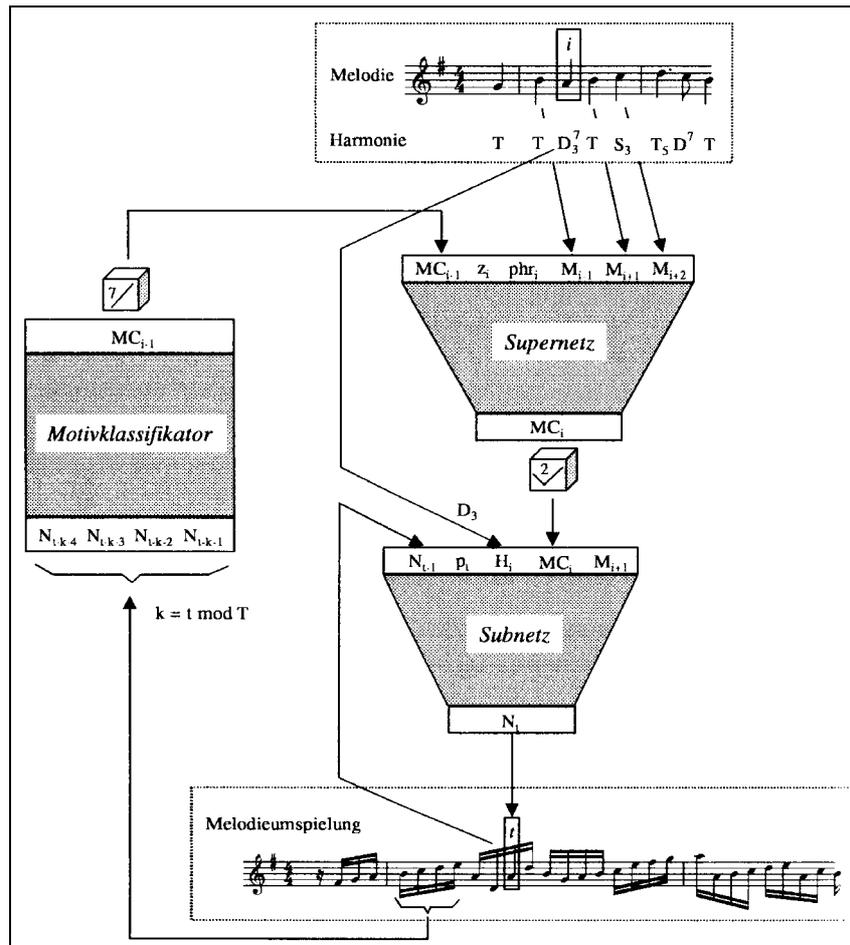
Mit Melonet sollen Chormelodien im Stil des Komponisten Johann Pachelbel um Melodieumspielungen ergänzt werden. Als Grundlage dienen 20 Choralabschnitte, von denen 12 als Trainingsmenge und 8 als Testmenge benutzt wurden.

#### **5.1.1.1 Aufbau**

Die Aufgabe von Melonet ist in zwei Schritte geteilt:

1. Die Harmonisierung der Melodie (auffinden von geeigneten Harmonien zu einer Melodie).
2. Das Versetzen der Melodie mit einer Melodieumspielung entsprechend der Harmonisierung.

Der erste Teil, die Harmonisierung, wird durch das später vorgestellte Harmonet durchgeführt. Daher wird im Folgenden davon ausgegangen, dass Melonet sich an einer vorhandenen Harmonisierung orientieren kann. Weiterhin wird die Eingabe auf ein festes Zeitraster von 16teln transformiert.

Abbildung 18: Struktur von Melonet<sup>30</sup>

Es stehen:

- $M_i$  für den aktuellen Melodieton (im Subnetz steht  $M_{i+1}$  für den nächsten Referenzton)
- $MC_i$  für die aktuelle Motivklasse
- $H_i$  für das Harmonische Feld
- $N_i$  für den aktuellen, zu bestimmenden Motivton der Melodieumspielung
- $z_i$  für die Position innerhalb des Taktes
- $phr_i$  für die Position innerhalb der Phrase
- $p_i$  für die Position innerhalb des Motivs

Vor dem Training werden die Umspielungen Pachelbels durch ein modifiziertes Kohonen-Netz gelernt und klassifiziert. Dieser dient schließlich als *Motivklassifikator*.

<sup>30</sup> Aus [Hörnel 00] S.122.

Für dieses Teilnetz wird die mögliche Anzahl von Klassen experimentell festgelegt (etwa mit 12 Klassen), da die Summe der verschiedenen Motive in allen Trainingsdaten zu groß war (über zweitausend).

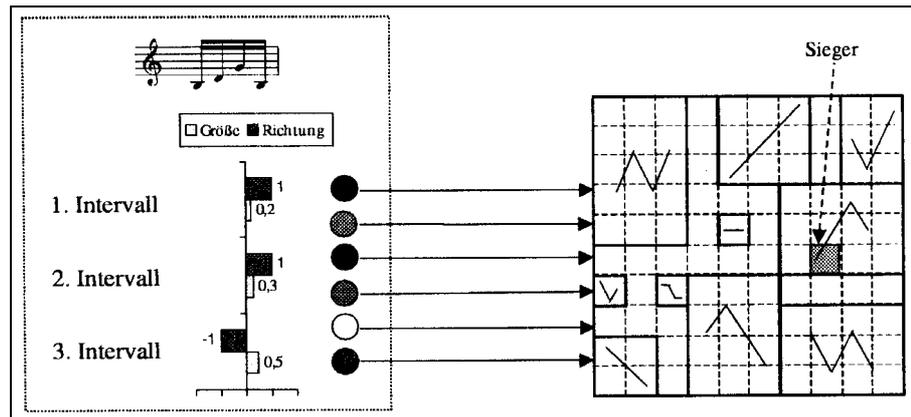


Abbildung 19: Kohonen-Netz aller Motive aus einer Choralumspielung<sup>31</sup>

Das Supernetz wird auf das Lernen von abstrakten Motivfolgen trainiert und soll entscheiden, welches Motiv zu einem Melodieton abhängig vom melodischen Kontext und den zuvor aufgetretenen Motiven am besten passt<sup>32</sup>.

Das Subnetz „lernt die Umsetzung der Motivklassen in konkrete Umspielungstöne abhängig vom gegebenen harmonischen Kontext.[...] Auch wenn die Ausgabe des Subnetzes hauptsächlich durch die Motivklasse [des Supernetzes] bestimmt ist, [...] muss es eine sinnvolle Umsetzung dieser Motivklasse bezüglich des harmonischen Kontexts finden. In manchen Fällen ist es möglich, dass das Subnetz ein Motiv bestimmt, das nicht zur Motivklasse des Supernetzes passt. Diese [vom Subnetz bestimmte] Motivklasse wird jedoch vom Supernetz berücksichtigt[...]“<sup>33</sup>

<sup>31</sup> Aus [Hörnel 00] S.126.

<sup>32</sup> [Hörnel 00] S.123.

<sup>33</sup> Aus [Hörnel 00] S.123/124.

### 5.1.1.2 Kodierung

Zur Kodierung der Töne wurde die nach Hörnel besonders effiziente, so genannte Intervallkodierung benutzt. Dabei wird jeder Ton durch seinen Abstand zum ersten Motivton (Referenzton) in der 1-von-n Form dargestellt.

Im Motiv vom Beethovens 5. Symphonie war das Motiv g-g-g-d. Der Referenzton wäre somit g. Der zweite Ton ist gleich, womit der Abstand 1 wäre (Prim, wir erinnern uns: bei Tonintervallen lautet der Abstand einer Note zu sich selbst eins). In der Kodierung würde dies dem Datensatz 010 0 1000000 entsprechen. Das Intervall zwischen dem dritten und dem vierten Ton (g-f-e-d) beträgt vier und ist damit eine absteigende Quarte. Die Kodierung lautet demnach 100 0 0000100.

Der besondere Vorteil dieser Kodierung ist die Unabhängigkeit der Darstellung von der gewählten Tonleiter. Unabhängig, ob das zu Grunde liegende Stück in der c-Dur oder a-Moll Tonleiter komponiert wurde, ist die Intervallkodierung in dieser Hinsicht allgemeingültig (transpositionsinvariant).

Zusätzlich kann ein Ergebnis in dieser Kodierung nur Noten aus der zu Grunde liegenden Tonart enthalten. Es kann also kein ganz fremder Ton gespielt werden (harmonisch invariant). Somit enthält bereits die Kodierung musikalisches Grundwissen.

	Richtung	Oktave	Intervallgröße	
<b>None</b>	↓	1 0 0	1	0 0 0 0 0 0 1
<b>Oktave</b>	↓	1 0 0	1	1 0 0 0 0 0 0
<b>Septime</b>	↓	1 0 0	0	0 1 0 0 0 0 0
<b>Sexte</b>	↓	1 0 0	0	0 0 1 0 0 0 0
<b>Quinte</b>	↓	1 0 0	0	0 0 0 1 0 0 0
<b>Quarte</b>	↓	1 0 0	0	0 0 0 0 1 0 0
<b>Terz</b>	↓	1 0 0	0	0 0 0 0 0 1 0
<b>Sekunde</b>	↓	1 0 0	0	0 0 0 0 0 0 1
<b>Prim</b>	→	0 1 0	0	1 0 0 0 0 0 0
<b>Sekunde</b>	↑	0 0 1	0	0 1 0 0 0 0 0
<b>Terz</b>	↑	0 0 1	0	0 0 1 0 0 0 0
<b>Quarte</b>	↑	0 0 1	0	0 0 0 1 0 0 0
<b>Quinte</b>	↑	0 0 1	0	0 0 0 0 1 0 0
<b>Sexte</b>	↑	0 0 1	0	0 0 0 0 0 1 0
<b>Septime</b>	↑	0 0 1	0	0 0 0 0 0 0 1
<b>Oktave</b>	↑	0 0 1	1	1 0 0 0 0 0 0
<b>None</b>	↑	0 0 1	1	0 1 0 0 0 0 0

Abbildung 20: Komplementäre Intervallkodierung<sup>34</sup>

<sup>34</sup> Aus [Hörnel 00] S.127.

### 5.1.1.3 Ergebnis

Abbildung 21: Ergebnisse der verschiedenen Netze zum Präludium V von J.S. Bach<sup>35</sup>

Um die Ergebnisse zu bewerten entwickelt Hörnel ein Fehlermaß für die Umspielungen. Fehler im Sinne einer Umspielung sind dabei: isolierte Töne (harmonieren nicht mit ihren Nachbartönen), fehlerhafte Sprünge (zwei Sprünge über mehr als Sext-Intervalle), repetierte Töne (Wiederholung des letzten Tones).

Das Netz mit 12 Motivklassen schneidet dabei mit einer Fehlerrate von etwa 7% am besten ab, gefolgt vom Netz mit 5 Motivklassen (8%) und dem Netz mit 20 Motivklassen (9,5%).

Als Verbesserung schlägt Hörnel vor, ein Komitee aus mehreren Netzen mit verschiedenen Anzahlen von Motivklassen zu benutzen und jeweils den besten Motivvorschlag auszuwählen.

Durch diese Verbesserung, so Hörnel, verringert sich der Fehler auf 0,9%, was einem unharmonischen Ton pro 100 Tönen entspricht.

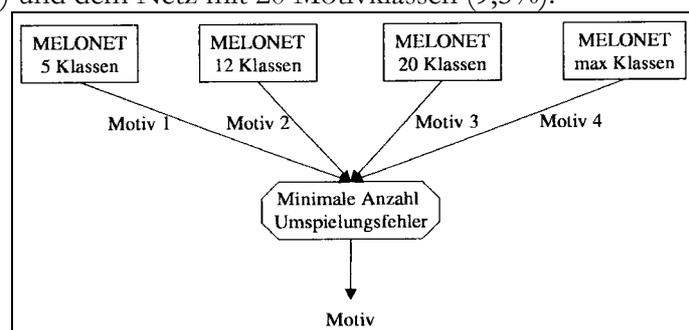


Abbildung 22: Komitee von mehreren Netzen<sup>36</sup>

<sup>35</sup> Aus [Hörnel 00] S.134.

<sup>36</sup> Aus [Hörnel 00] S.136.

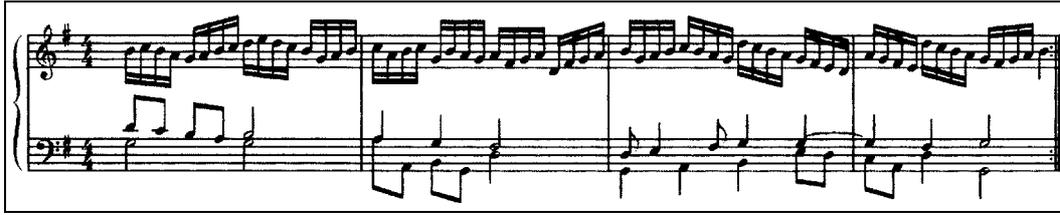


Abbildung 23: Ausschnitt einer Melodievariation über die Chormelodie „Werde munter, mein Gemüte“ vom Melonet-Komitee<sup>37</sup>

## 5.1.2 Jazznet

Mit Jazznet sollen Jazz-Soloimprovisationen im Stile des legendären Saxophonisten Charlie Parker<sup>38</sup>, der als Urvater des Bebop gilt, erzeugt werden.

### 5.1.2.1 Aufbau

Jazznet basiert auf Melonet, wurde aber den folgenden neuen Anforderungen angepasst:<sup>39</sup>

- Es gibt keine Melodievorgabe mehr. Die Improvisation basiert nur auf den Harmonien, die durch die Akkorde vorgegeben werden.
- Es sind Rhythmisch feinere Umspielungen möglich. Im Gegensatz zu den sechzehntel Noten liegen im Jazz achte Noten mit sechzehntel Noten und achte Triolen als Variation. Hierfür werden achte Triolen als Abfolge einer achte und zwei sechzehntel Noten modelliert.
- Um eine größere Freiheit der Improvisation zu erreichen, soll das System eine stochastische Komponente enthalten, damit über gleiche Akkordfolgen unterschiedlich improvisiert wird.

Es wird eine rhythmische Verfeinerung derart eingeführt, dass nun nicht ganze Motive, sondern Teilmotive benutzt werden. In Jazz, das bewegter ist als Choräle, ist dies üblich. Erneut wird das Teilnetz zur Klassifikation der Teilmotive mit den Vorgaben trainiert.

Weiterhin wird ein Stück wie folgt improvisiert:

<sup>37</sup> Aus [Hörnel 00] S.138.

<sup>38</sup> Charlie Parker, amerik. Jazzsaxophonist und Godfather des Bebop Jazzstils (1920-1955).

<sup>39</sup> [Hörnel 00] S.140.

Dem Netz wird eine Teilphrase (ein Anfang) gegeben. Das Referenzton-Netz sagt den nächsten Referenzton (erster Ton des nächsten Motivs) voraus. Der Teilmotiv Klassifikator klassifiziert die Teilphrase. Aus diesen Informationen wird von einem weiteren Netz eine Improvisationslinie erzeugt (Vervollständigung der Noten bis zum nächsten Referenzton). Schließlich werden Verzierungen der Improvisationslinie erzeugt.

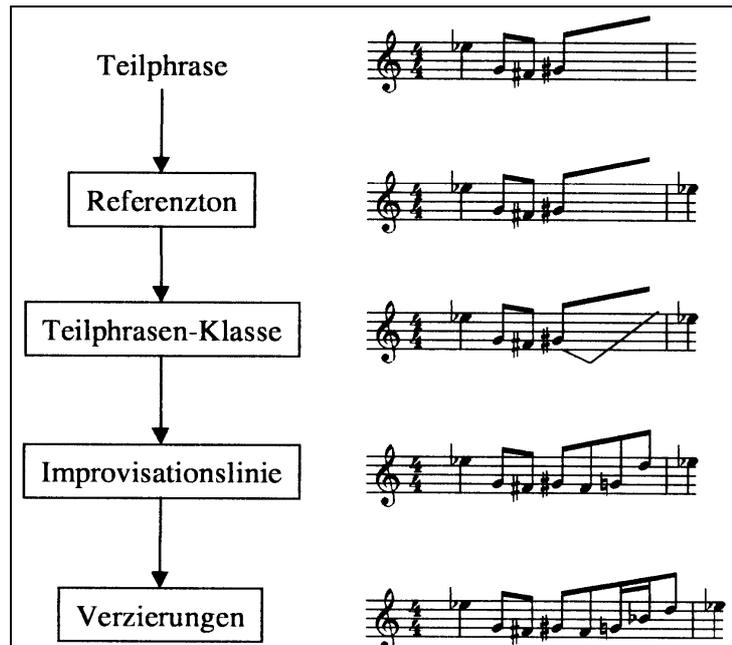


Abbildung 24: Hierarchische Bestimmung einer Teilphrase durch Jazznet<sup>40</sup>

### 5.1.2.2 Ergebnis

Um das Ergebnis von Jazznet auszuwerten, vergleicht Hörnel Jazznet mit einem regelbasierten System namens JazzRule, das durch statistische Auswertungen der Parker Stücke gewonnen wurde, sowie mit den Originalen von Parker.

Als Vergleichsmerkmale dienten dabei:

- Verlauf der Melodiekontur
- Harmonische Beziehung zwischen Melodietönen und darunter liegenden Akkorden
- Geeignete Verwendung von Zieltönen
- Auftreten Parker-ähnlicher Muster

Es stellte sich heraus, dass Parker hauptsächlich in fallender Richtung improvisierte, also die Töne im Durchschnitt von höheren Tönen zu tieferen spielte. Diese Eigenschaft wurde von Jazznet besser als von Jazzrule erfasst. Hingegen waren die lokalen Schwankungen von Jazzrule näher an den Originalen von Parker.

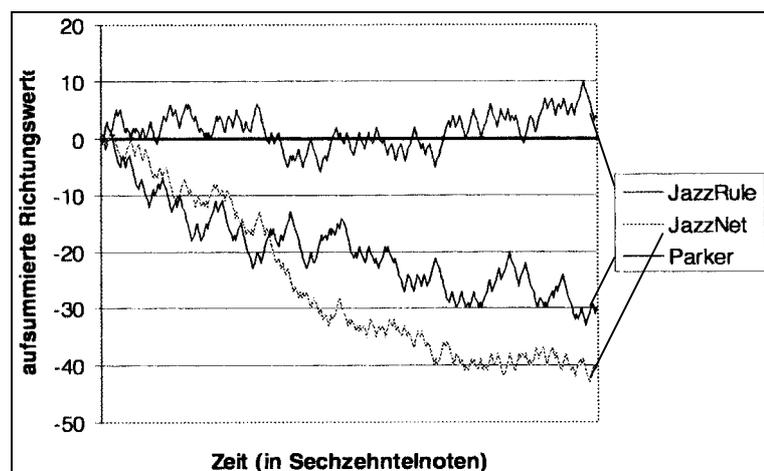


Abbildung 25: Aufsummierte Richtungswerte für Jazzrule, Jazznet und Parker<sup>41</sup>

<sup>40</sup> Aus [Hörnel 00] S.141.

Bei der Analyse der harmonischen Beziehung zwischen Melodie und darunter liegenden Harmonien wurden alle Töne einer von drei Klassen zugeordnet:<sup>42</sup>

- inside (harmonieeigen), wenn der Ton im Akkord auftaucht,
- leading (Leitton), wenn der Ton einen Halbton unter einem harmonieeigenen Ton liegt,
- outside (harmoniefremd), sonst

Auch diese Analyse ergab, das Jazznet dem Parker Original näher liegt, während Jazzrule zu viele Outside-Töne produzierte.

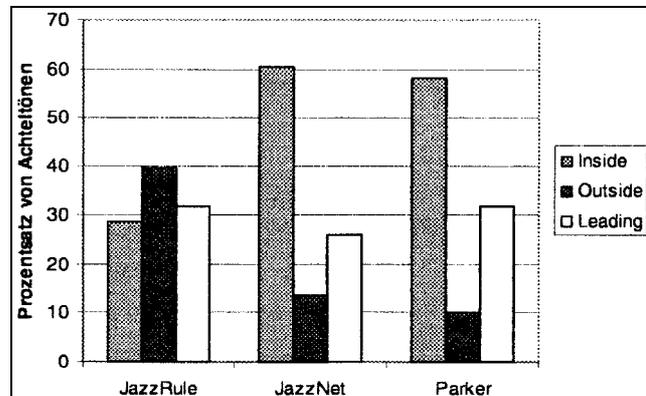


Abbildung 26: Verteilung der Inside, Outside und Leading Achteltöne<sup>43</sup>



Abbildung 27: Beispielimprovisationen von Jazzrule, Jazznet und Charlie Parker<sup>44</sup>

## 5.2 Melodiekomposition – Melogenet

Im Folgenden sollen die Netze eine Kompositionstätigkeit ausüben und unterliegen daher nicht dem einfachen linken Kontext der Improvisation. Zu jeder Zeit kann eine bereits festgelegte Note revidiert werden, was bei der Improvisation nicht möglich war. Daher verwendet Hörnel in den folgenden Modellen genetische Algorithmen, um die gesamte Melodie zu Beginn festzulegen und im Verlauf der Evolutionen zu verbessern. Dies entspricht auch dem tatsächlichen Vorgehen von z.B. Beethoven, der zuerst einige Motive als Grundidee notierte um diese dann durch mehrere Ergänzungen und Veränderungen zu einem fertigen Thema zu überführen.<sup>45</sup>

„Die Betrachtung des Kompositionsvorganges als evolutionären Prozess bildet den Ausgangspunkt der beiden Systeme Melogenet und Harmogenet [...]. Zur Entwicklung lokaler Abschnitte einer Melodie bzw. einer Harmonisierung werden evolutionäre Algorithmen verwendet [...]. Die Bewertungs- oder *Fitnessfunktion*, durch welche die Evolution gesteuert wird,

<sup>41</sup> Intervall zweier benachbarter Töne: +1 für aufsteigend, -1 für absteigend, 0 für konstant. Die Werte wurden summiert. Aus [Hörnel 00] S.143.

<sup>42</sup> [Hörnel 00] S.144.

<sup>43</sup> Aus [Hörnel 00] S.144.

<sup>44</sup> Aus [Hörnel 00] S.146.

<sup>45</sup> [Hörnel 00] S.147.



## 5.2.1 Melogenet

### 5.2.1.1 Aufbau

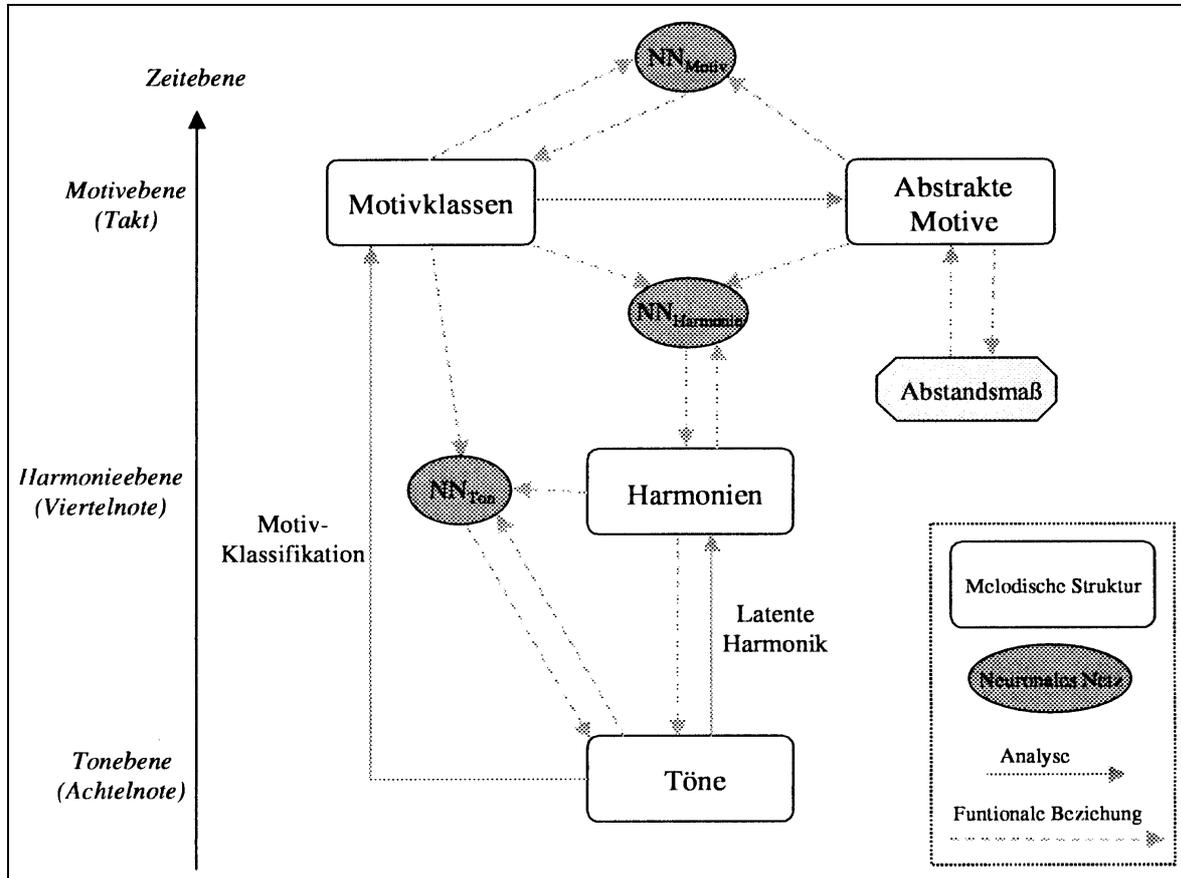


Abbildung 29: Analyse und funktionale Abhängigkeiten in Melogenet<sup>49</sup>

Melogenet besteht aus mehreren unabhängigen Teilnetzen und einem Netz, das die Ausgaben der Teilnetze verarbeitet.

Das Teilnetz  $NN_{Motiv}$  lernt eine Motivklasse in Abhängigkeit von benachbarten Motivklassen und Motivfolgen zu lernen. Dabei macht es gebrauch von einem Motivklassifikator.

Das Teilnetz  $NN_{Harmonie}$  lernt eine Harmonie aus benachbarten Harmonien, der Motivklasse und der Motivfolge.

Das Teilnetz  $NN_{Referenzton}$  und  $NN_{Motivton}$  bestimmen die Töne aus benachbarten Tönen, der Harmonie und der Motivklasse.

Dabei wird, im Vergleich zu den Improvisationsmodellen, auch der rechte Kontext ausgewertet. Nachdem die Netze trainiert wurden, wird der evolutionäre Algorithmus verwendet, um die Fitnessfunktion zu optimieren, die aus dem gelernten Wissen der verschiedenen Netze zusammengesetzt wird. „Zur Bestimmung der Fitness wird zu einer gegebenen Melodie für jedes neuronale Netz die Differenz zwischen den durch das Netz vorhergesagten und den tatsächlich vorhandenen Elementen berechnet. Je kleiner diese Differenz ist, umso höher die Fitness der Melodie.“<sup>50</sup>

<sup>49</sup> Aus [Hörnel 00] S.153.

<sup>50</sup> Aus [Hörnel 00] S.159.

### 5.2.1.2 Ergebnis

Getestet wurde Melogenet mit 32 Kinderliedern aus der Essener Liederdatenbank. Zur Vervollständigung wurden jeweils zwei Anfangstakte vorgegeben. Nach 100 Generationen wurde die Evolution abgebrochen. Es zeigte sich, dass der Fitnesswert bis zu einem Wert von etwa 0,9 stetig verbesserte, danach aber trotz Fitnessverbesserung zu glatteren, weniger kreativen Melodien konvergierte. Außerdem setzten sich die Motive, die in der Trainingsmenge am häufigsten vorkamen, durch. Durch die Ergänzung der Fitnessfunktion um eine Bewertung für die Motivabwechslungen konnte das Netz jedoch verbessert werden.

Das gleiche Experiment wurde mit 41 chinesischen Volksliedern im Shanxi-Stil durchgeführt, um die Fähigkeiten des Systems auf anderen Stilen zu testen. Die Ergebnisse waren vergleichbar, auch wenn die Fitness im Durchschnitt etwas unter der der Kinderlieder lag.

Schließlich wurde Melogenet zum bestimmen eines Stiles eingesetzt, indem Testlieder dem Shanxi-System und dem Kinderlieder-System gegeben wurden. Das Netz, das eine bessere Fitness lieferte, gab mit höherer Wahrscheinlichkeit an, dass die Eingabe zu seinem Stil passte.

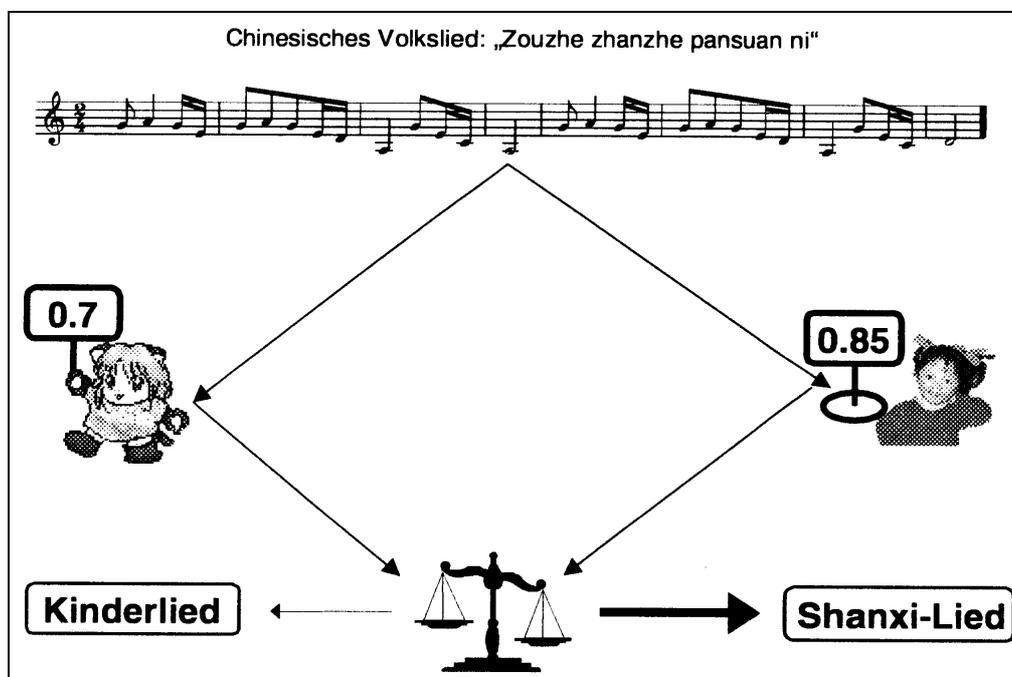


Abbildung 30: Stilerkennung eines chinesischen Volksliedes durch automatische Fitneßbewertung<sup>51</sup>

Nach diesem Konzept kann ein Komitee von Experten verschiedener Stile dazu genutzt werden, ein Lied einem bestimmten Stil zuzuordnen.

In einem abschließenden Versuch wurden Probanden künstliche, durch Melogenet generierte und Originallieder präsentiert. Die Probanden sollten die Originale von den synthetischen Liedern unterscheiden. Das Experiment ergab, dass die Probanden diejenigen Lieder als Originale identifizierten, die ihnen auch besser gefielen. Melogenet hatte relevante Stilmerkmale gelernt, es gab aber weitere wichtige melodische Aspekte, die nicht erfasst wurden.

<sup>51</sup> Aus [Hörnel 00] S.167.

Abbildung 31: Melodien, die für das Hörexperiment verwendet wurden und deren Fitness<sup>52</sup>

	Mel. 1	Mel. 2	Mel. 3	Mel. 4	Mel. 5	Mel. 6	Mel. 7	Mel. 8	Mel. 9	Mel. 10
<b>Teiln. 1</b>	o		o	o			o			o
<b>Teiln. 2</b>			o		o			o	o	o
<b>Teiln. 3</b>			o		o		o		o	o
<b>Teiln. 4</b>		o	o				o			o
<b>Teiln. 5</b>	o		o	o			o		o	o
<b>Teiln. 6</b>		o			o		o		o	o
<b>Teiln. 7</b>		o			o				o	o
<b>Teiln. 8</b>	o	o	o					o	o	o
<b>Teiln. 9</b>		o		o	o		o		o	o
<b>Teiln. 10</b>		o			o		o		o	o
<b>Teiln. 11</b>				o	o		o	o	o	o
<b>korrekt</b>		o			o		o		o	o

Abbildung 32: Ergebnis des Experiments, aufgeteilt nach Teilnehmer (o=Teilnehmer tippt auf Original)<sup>53</sup>

<sup>52</sup> Aus [Hörnel 00] S.173.

<sup>53</sup> Aus [Hörnel 00] S.174.

### 5.3 Choralharmonisierung – Harmonet und Harmogenet

Bei der Choralharmonisierung gilt es für eine gegebene Melodie eine Harmoniefolge zu definieren, so dass begleitende Akkorde zur Melodie entstehen.

Auch hier gibt es regelbasierte Systeme, wie z.B. CHORAL, das mit über 500 Regeln im Stile Bachs harmonisiert. Diesem System liegt jedoch ein unheimlicher zeitlicher und intellektueller Aufwand zugrunde, und dabei ist das System auf einen konkreten Stil festgelegt.

#### 5.3.1 Harmonet

Harmonet wurde von H. Hilden entwickelt um im Stil Bachs zu harmonisieren und von Hörnel erweitert, um beliebige Stile zu lernen.

##### 5.3.1.1 Aufbau

Die Harmonisierung durch Harmonet wird in drei Schritten vollbracht. Im ersten Schritt wird anhand der Chormelodie und evtl. bereits festgelegten Harmonien die nächste Harmonie festgelegt. Im zweiten Schritt wird der ermittelten Harmonie ein Akkord zugewiesen. Im letzten Schritt werden Achtelumspielungen hinzugefügt.

Im ersten Schritt werden mehrere Netze eingesetzt, um aus dem Kontext der letzten drei vorangegangenen Harmonien, dem letzten, dem aktuellen und dem nächsten Melodieton sowie der Position in der Phrase die aktuelle Harmonie zu bestimmen. Dieser Teil der Aufgabe ist der schwierigste. Daher werden die Kriterien für die Bestimmung einer Harmonie aufgeteilt.

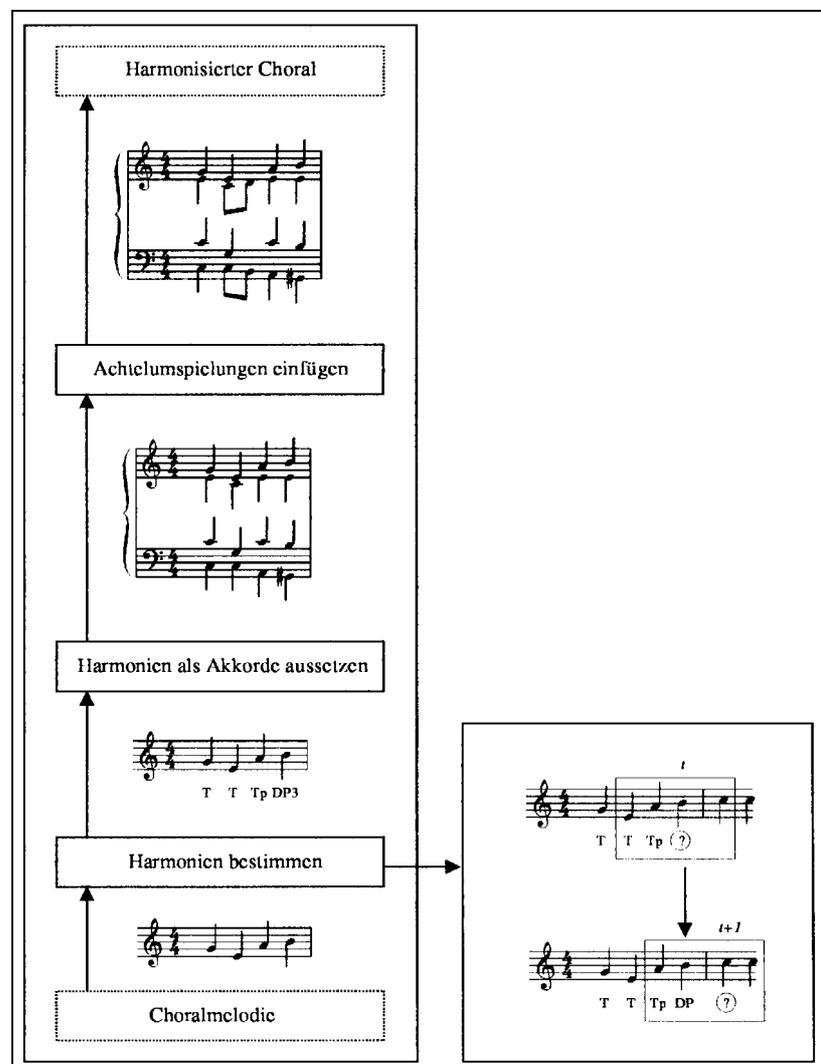


Abbildung 33: Aufbau von Harmonet<sup>54</sup>

<sup>54</sup> Aus [Hörnel 00] S.181.

Harmonien werden hier in drei Teilen ermittelt:

1. Harmonische Funktion (Tonikas, Subdominanten und Dominanten bezüglich einer Tonart). Hierzu werden drei Netze eingesetzt, die jeweils mit verschiedenen Kontextlängen arbeiten. Die Töne werden über eine harmonische Kodierung kodiert, die angibt, in welchen von 12 harmonischen Funktionen der Ton vorkommt. Das Ergebnis ist die aktuelle Harmonie.
2. Harmonische Umkehrung (Stellung des Basstones im Akkord). Ein weiteres Netz bestimmt anhand der Kontextinformation sowie der ermittelten Harmonie die Lage des Basstones.
3. Charakteristische Dissonanzen (ergänzte Sexten, Septimen). Ein letztes Netz fügt Töne so hinzu, dass bestimmte Akkorde, die durch die Harmonischen Funktionen nicht berücksichtigt wurden aber durch ergänzen einer Note hergestellt werden können.

	T	D	S	Tp	Sp	Dp	DD	DP	TP	d	VTp	SS
<b>c</b>	1	0	1	1	0	0	1	0	0	0	0	0
<b>cis/des</b>	0	0	0	0	0	0	0	0	1	0	0	0
<b>d</b>	0	1	0	0	1	0	1	1	0	1	1	1
...												

Abbildung 34: Harmonische Tonkodierung für c-Dur<sup>55</sup>

Im zweiten Schritt wird ein symbolischer Algorithmus (basierend auf Regeln) eingesetzt, um gemäß der Stimmführungsregeln und der zugrunde liegenden Harmonie einen Akkord auszuwählen. „Der fließende Verlauf der Stimmen wird durch globale Optimierung der Stimmführung unter Verwendung dynamischen Programmierens [...] gewährleistet.“<sup>56</sup>

Im dritten Schritt wird ein Netz trainiert, das die Umspielungen anhand der Harmonien und Melodietöne festlegt.

### 5.3.1.2 Ergebnis

„Harmonet produziert in vielen Fällen interessante und stilechte Harmonisierungen und findet auch für Nicht-Choralmelodien oft überzeugende Lösungen [...]. Trotzdem zeigen sich einige Schwächen[...]:

- Fehlerhafte/Langweilige Harmonisierung bei zu kleiner/großer Trainingsmenge
- Probleme der Bassführung durch den sequentiellen Harmonisierungsprozess
- Stilunabhängige Aussetzung
- Tonale Zentren werden nicht modelliert“<sup>57</sup>

Einige dieser Probleme werden in Harmogenet gelöst.

<sup>55</sup> Aus [Hörnel 00] S.185.

<sup>56</sup> Aus [Hörnel 00] S.182.

<sup>57</sup> Aus [Hörnel 00] S.186.

## 5.3.2 Harmogenet

### 5.3.2.1 Aufbau

Wie bei Harmonet wird die Harmonisierung in drei Schritten durchgeführt. Lediglich Schritt eins wird geändert. Hier werden wieder genetische Algorithmen eingesetzt, um die aktuelle Harmonie zu bestimmen. Gleichzeitig werden die Netze aus der Harmonet Harmonisierung so erweitert, dass sie nun auch den rechten Kontext auswerten können. Durch ein weiteres Netz wird das tonale Zentrum ermittelt. Zwei weitere Netze kontrollieren die Aussetzungen ( $N_7$  und  $N_8$ ). Schließlich wird für die Bassführung ebenfalls ein eigenes Netz benutzt ( $N_9$ ). Die Fitnessbewertung ist äquivalent zu Melogenet.

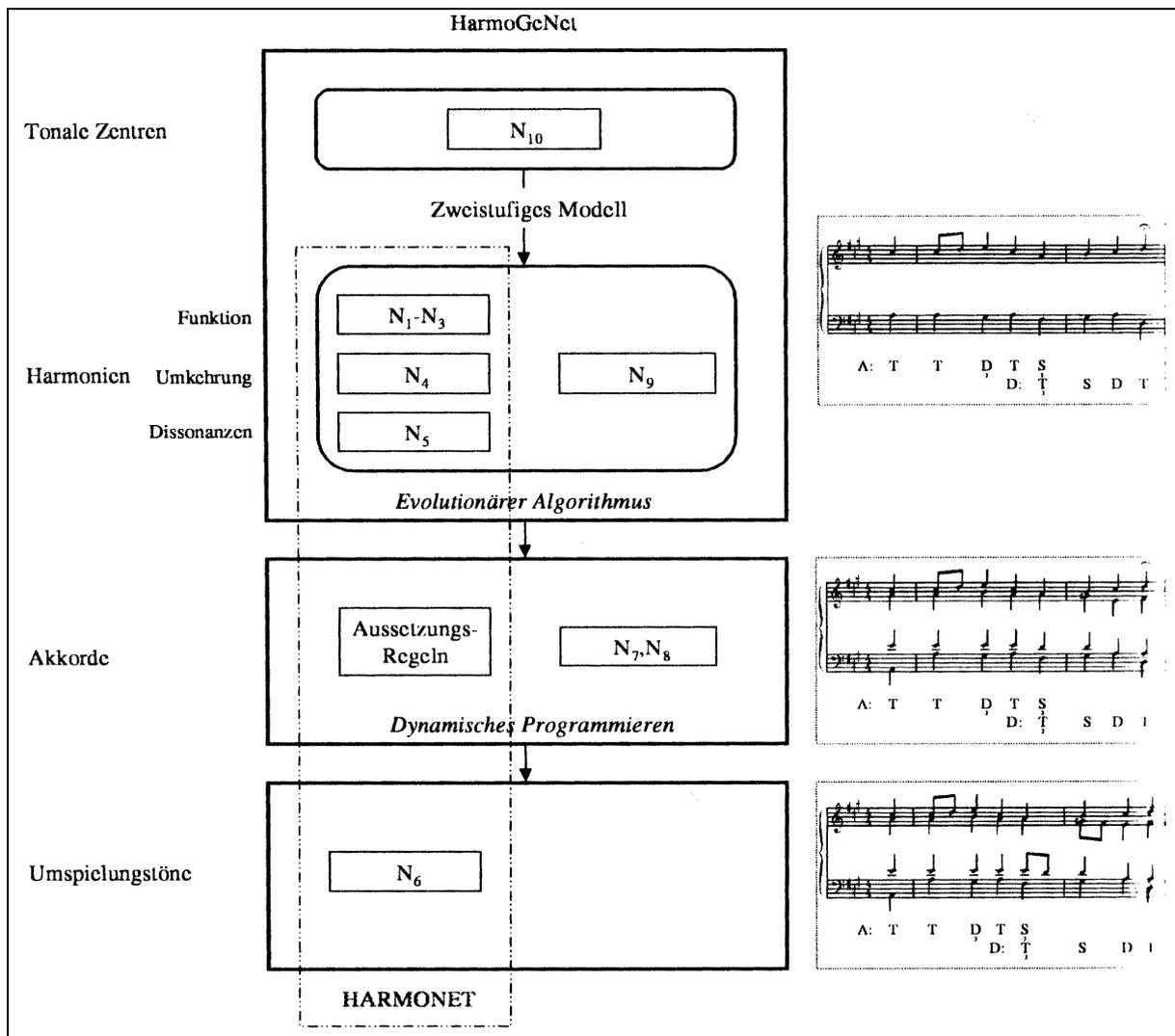


Abbildung 35: Aufbau von Harmogenet<sup>58</sup>

<sup>58</sup> Aus [Hörnel 00] S.188.

### 5.3.2.2 Ergebnis

Mit Hilfe der tonalen Zentren und dem rechten Kontext wird die Klassifikationsgüte des Systems auf den Testdaten um 10% auf 76,9% erhöht.

## 5.4 Stilanalyse und Stilerkennung

Wie bereits bei den Shanxi- und den Kinderliedern besteht die Stilerkennung aus einem Komitee von Stilexperten, die eine Klassifikation auf einem unbekanntem Lied durchführen. Derjenige Stilexperte, der die höchste Klassifikationsgüte erreicht hat den Stil identifiziert. Dies wird auch *vergleichende Stilanalyse* genannt.

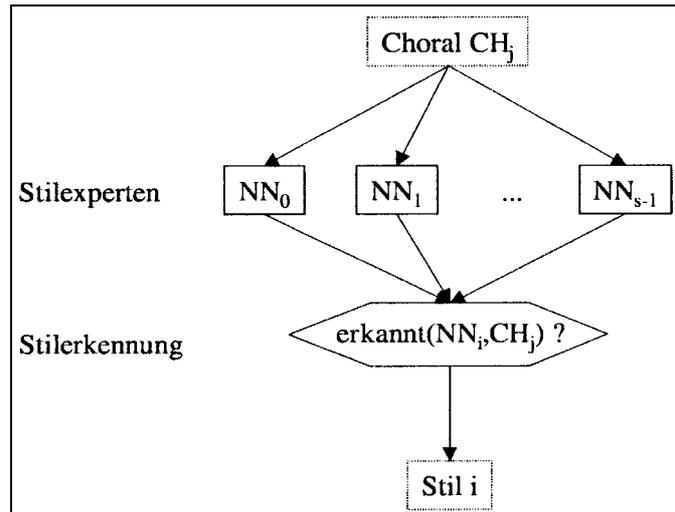


Abbildung 36: Stilerkennungsprozess<sup>59</sup>

(Die Stilerkennungsrate des Bachnetzes auf Reger-Choräle ist auf die kleine Reger-Trainingsmenge sowie die Inhomogenität der Reger-Choräle im Vergleich zu denen von Bach oder Scheidt zurück zu führen).

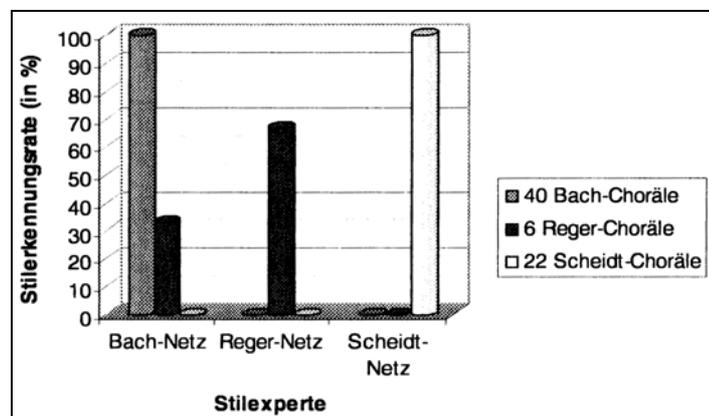


Abbildung 37: Stilerkennungsgüte für drei Netze (Bach-, Reger<sup>60</sup>- und Scheidt<sup>61</sup>-Stil)<sup>62</sup>

<sup>59</sup> Aus [Hörnel 00] S.204.

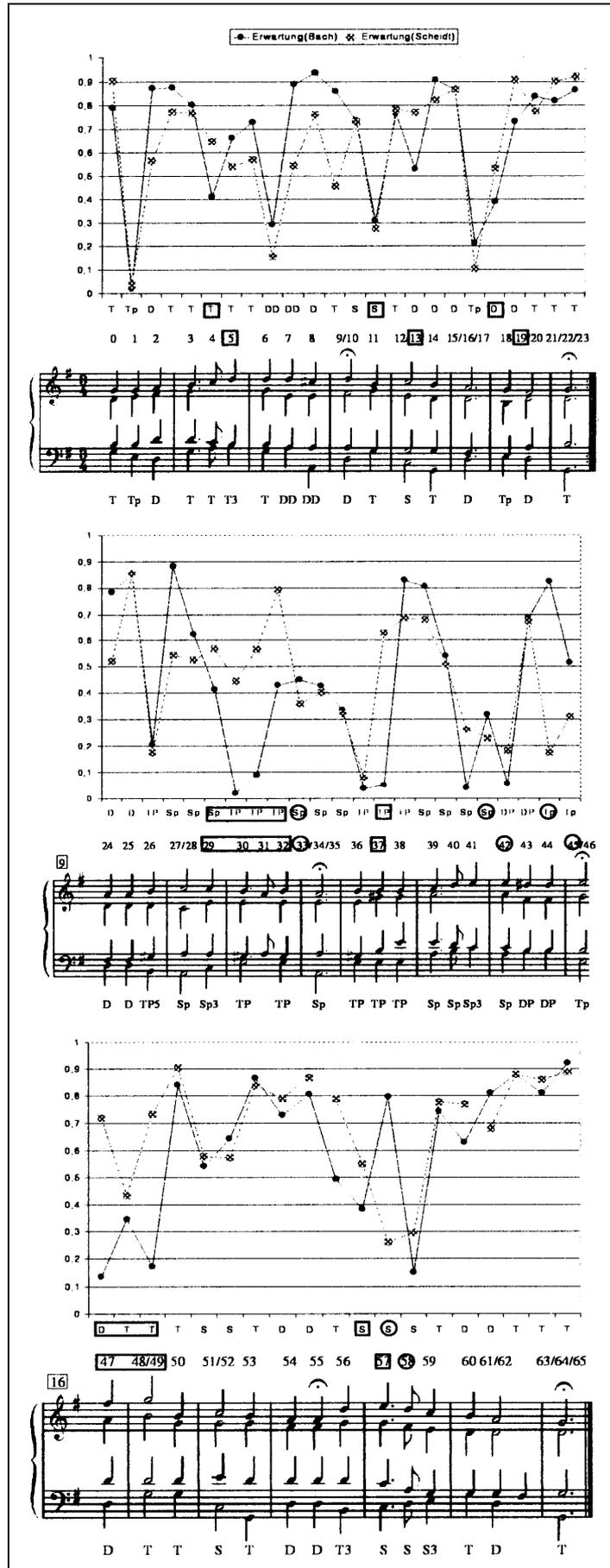
<sup>60</sup> Max Reger, dt. Komponist (1873-1912).

<sup>61</sup> Samuel Scheidt, dt. Komponist (1587-1654).

<sup>62</sup> Aus [Hörnel 00] S.208.

Bei den Tests fällt Hörnel eine Anomalie auf: Im Bach-Choral „Du Lebensfürst, Herr Christ“ ist eine Passage, die von den Stilexperten eher Scheidt als Bach zugeordnet wird. Nach Nachforschungen stellt Hörnel fest, dass es bereits fachliche Arbeiten gibt, die darauf hinweisen, dass diese Passage Bach unspezifisch ist und dass vermutet wird, Bach habe diese Passage aus unbekanntem Gründen eins zu eins von Scheidt übernommen. Einen wirklichen Beweis für diese Behauptung gibt es zwar nicht, aber der Fall demonstriert die Güte der Systeme.

Abbildung 38: Harmonische Erwartungen des Bach- und Scheidt-Netzes für den Choral<sup>63</sup>



<sup>63</sup> Aus [Hörnel 00] S.221.

## 6 Literaturverzeichnis

- [Hörnel 00] Hörnel, Dominik: „Lernen musikalischer Strukturen und Stile mit neuronalen Netzen“  
Dissertation zur Erlangung des akademischen Doktorgrades.  
Shaker Verlag, 2000
- [Ziegenrucker 97] Ziegenrucker, Wieland: „ABC Musik – Allgemeine Musiklehre“  
Verlag Breitkopf & Härtel, 1997
- [www.classicalarchives.com/beethovn.html](http://www.classicalarchives.com/beethovn.html)  
Internetseite über Beethoven.
- [dict.leo.org](http://dict.leo.org)  
Internetseite mit Englisch<->Deutsch Übersetzer.
- [www.worldjazz.ch](http://www.worldjazz.ch)  
Internetseiten über Jazzmusik.