© Naturwiss.-med. Ver. Innsbruck; download unter www.biologiezentrum.at

Ber. nat.-med. Verein Innsbruck	Band 79	S. 57 – 61	Innsbruck, Okt. 1992

ROT3D & COAGUL4D - 2 Programme zur Auswertung von Ordinationen *)

von

Michael HAUSER **)

ROT3D & COAGUL4D - 2 Programs for Easy Ordination-Analysis

Abstracts: Although already common in the businessworld there is still much to be done in the scientific world regarding the use of multivariate and statistical programs. The programs here introduced are the first two of a series of fast and easy-to-use applications written for the IBM-PC/AT/386(TM) (and compatibles). Grafic screen, full Mouse support and a "GEM(TM)-like"-User interface are the fundamental principles employed.

ROT3D (Version 1.4) (Fig. 1) allows real-time 3-dimensional Rotation of the scores of 3 Ordination-axes on screen. Besides many Mark- and Hide-options datapoints can be identified by simply clicking at them.

The datapoints can be connected, so that they represent clusters (or groups). It is therefore possible either to underlay an already existing classification or to classify the ordination, e.g. with COAGUL4D (Version 1.0) (Fig. 2), that produces as output the clusters at a chosen distance-level using Nearest Neighbour algorithm and Euclidian Distance (4-dimensional). Basically an incomplete Minimum Spanning Tree of the Ordination scores is obtained.

It is planned to incorporate the next versions of these programs in a complete Multivariate-Analysis and Pattern-Analysis Package available not only for the PC but also for the Apple Macintosh(TM) and the Atari ST(TM). In fact parts of or even whole applications should already be available. For further information, please contact the author.

Keywords: Klassifikation, Grafik, Identifikation, Ordination, UserInterface.

1. Einleitung:

Als zu Beginn dieses Jahrzehnts erstmals Personal Computer (PC's) auch preislich verfügbar waren, leiteten sie unter anderem auch im wissenschaftlichen Bereich eine vollkommen neue Arbeitsweise ein. War man früher Großrechnern (und deren Personal) auf Gedeih und Verderb ausgeliefert, so ist es heutzutage möglich, die meisten Probleme, auch die der multivariaten und statistischen Analyse, auf dem eigenen Desk- oder Laptop in erträglicher Zeit zu lösen. Der IBM-PC(TM) (und mehr noch: alle seine Clones und Nachfolger) hat zum allergrößten Teil diesen "Quantensprung" möglich gemacht. Der Einfluß dieser Transition hat aber diesem Typ Computer gerade den deutlichsten Stempel aufgedrückt: Die meisten analytischen Programme sind nichts anderes als Portierungen von Großrechnerversionen, die sich genauso unbarmherzig Eingabefehlern gegenüber zeigen und keinerlei Hilfestellung geben.

*) Vortrag, gehalten im Rahmen des 5. Österreichischen Botanikertreffens in Innsbruck, 25. - 28. Mai 1989.

**) Anschrift des Verfassers: M. Hauser, Institut für Pflanzenphysiologie der Universität Wien, Althanstraße 14, Postfach 285, A-1091 Wien, Österreich.

© Naturwiss.-med. Ver. Innsbruck; download unter www.biologiezentrum.at

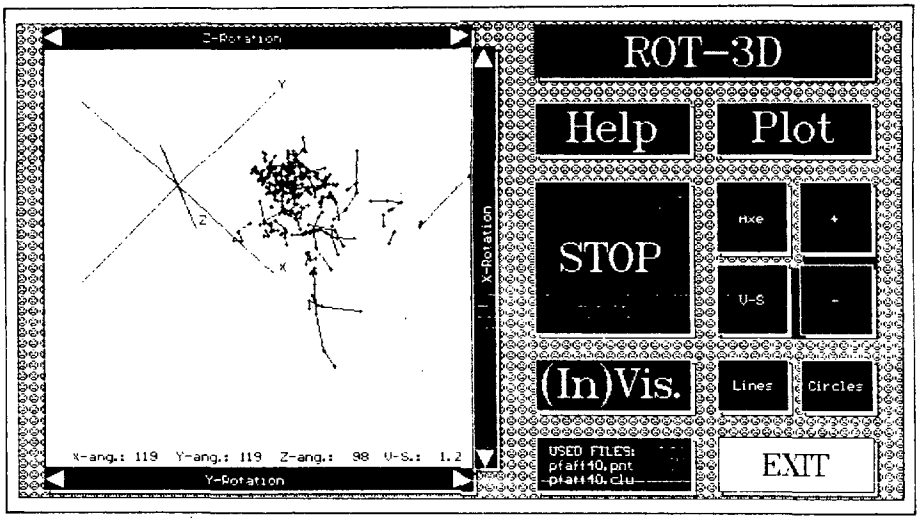

Abb. 1: Desktop von ROT3D.

Selbst als um die Mitte dieses Jahrzehnts Personal Computer der nächsten Generation mit grafischer Oberfläche, Mausbedienung, Menüleisten und Fenstern erschienen, hatten sie lange Zeit keinen Einfluß auf das Aussehen vor allem wissenschaftlicher Programme. Diese Ignoranz wurde von Programmierern zum Teil hinter dem Mantel der Portabilität, zum anderen durch Zeitmangel und Programmieraufwand gerechtfertigt. Sie vergaßen aber dabei 3 grundlegende Prinzipien:

1) Programme sollen den Bedürfnissen des Users entsprechen und nicht der User den Bedürfnissen und Fähigkeiten des Programmierers.

2) als visuelles Geschöpf erkennt der Mensch Zusammenhänge viel eher in Form von Grafiken als in und zwischen den Zeilen meterlanger Tabellen, und

3) gerade beginnende Computeruser werden durch schlechte oder schlecht gemachte Software eher abgeschreckt als ermutigt. Wie viele Erkenntnisse und Zusammenhänge mögen wohl nicht gewonnen oder erkannt worden sein, nur weil die Programmumgebung ein spielerisches (und damit entspanntes) Arbeiten mit dem Datenmaterial nicht ermöglicht hat? (siehe hiezu GLEICK, 1987, NAIMAN et al., 1988, aber auch SOUČEK & SOUČEK, 1988).

Der Grund der Portabilität wird hinfällig, wenn die meisten Datensätze am eigenen PC bearbeitet werden können und werden, der erhöhte Programmieraufwand wird sowohl durch die Zuhilfenahme geeigneter Tools (Programmierhilfen), als auch durch geeignete Programmstrukturierung (und auch die Programmierung in anderen Sprachen als FORTRAN) erheblich minimiert.

Sowohl diese Überlegungen als auch eigene Bedürfnisse haben zur Entwicklung eines Programmpaketes geführt, wovon die ersten (auch unabhängig lauffähigen) Applikationen hier beschrieben werden.

2. Programmbeschreibungen:

Es kann hier nur eine grobe Beschreibung der Details, wie Datenformat, geben. Nähere Informationen sind auf jeder der Programmdisketten als ASCII-File abgespeichert.

ROT3D ist ein Echtzeit-3D-Rotationsprogramm, das dem User erlaubt, beliebige Ordinationskoordinatendaten der ersten 3 Komponentenachsen um 3 Bildschirmachsen – auch gleichzei-

© Naturwiss.-med. Ver. Innsbruck; download unter www.biologiezentrum.at

tig – zu rotieren, wobei die ersten beiden Rotationsachsen waagrecht bzw. senkrecht in der Bildschirmebene liegen, die dritte orthogonal zu beiden anderen steht.

Das Programm verarbeitet in der momentanen Version die Koordinaten von 1000 Objekten und 1000 Clustern (siehe weiter unten), also mehr, als die meisten Ordinationsprogramme bearbeiten können. Die Größe des .EXE-Files beträgt 110.000 Bytes.

Als Eingabefiles dienen:

1) ein Koordinatenfile (.PNT), das neben den Koordinatenwerten für 4 Achsen (momentan werden nur die ersten 3 ausgewertet, das Format ist aber im Sinne der Aufwärtskompatibilität so gewählt) auch noch zu jedem Objekt einen Kommentar erhalten kann (z.B. der Artname, wenn die Objekte Arten sind),

2) ein Verbindungsfile (.CLU), das es erlaubt, mittels einer bereits bestehenden Klassifikation die Objekte zu Gruppen zusammenzufassen. Dieses File erlaubt es – ähnlich dem Koordinatenfile – für jede Gruppe (Cluster) einen Kommentar anzuführen (z.B. wenn die Objekte Arten sind, kann hier die jeweilige Gruppenbezeichnung für ökologische Parameter stehen).

Nachdem das Programm die beiden Datenfiles eingelesen hat, erscheint am Monitor das Desktop (Abb. 1). Es ist in 2 Bereiche unterteilt, links das eigentliche Betrachtungsfenster für die Daten, das von Schiebern für die jeweiligen Rotationen eingerahmt wird, rechts das Steuerpult mit sogenannten Buttons (also Knöpfen), die bestimmte Aktionen auslösen:

Durch Anklicken mit der Maus werden bei einigen Fenster eröffnet, die zusätzliche Information tragen oder weitere Aktionen ermöglichen. So liefert HELP eine komplette Befehlsliste, PLOT eröffnet die Möglichkeit, das Beobachtungsfenster auf einem EPSON-FX(TM) (oder kompatiblen) Drucker ausgeben zu können (in der nächsten Version wird zusätzlich auch Ausgabe auf Plotter möglich sein). (IN)VIS erlaubt es, in einem eigenen Fenster Gruppen (un)sichtbar zu machen. Neben verschiedenen Darstellungsoptionen (CIRCLES, LINES, AXE, V-S, +,–) eröffnet bei geSTOPter Darstellung das Anklicken eines Datenpunktes ein eigenes Fenster, das neben den Koordinatenwerten des Punktes auch noch die jeweiligen Kommentare enthält.

Oftmals kann es von Vorteil sein, eine bereits bestehende Ordination nachträglich zu klassifizieren. Das Programm

COAGUL4D ermöglicht eine flexible Klassifikation von Ordinationen (bis zu 4 Dimensionen), wobei die Gruppenweite vom User im Bereich zwischen kleinster und größter Nearest Neighbour-Euklidischer Distanz (NNED) eingestellt werden kann. Das Programm erzeugt einen unvollständigen Minimum-Spanning-Tree (MST), bei dem die einzelnen Cluster 1. Ordnung nicht weiter verknüpft werden (die nächste Version wird eine komplette MST-Analysis beinhalten). Näheres dazu unter anderem bei ORLCI (1978) und PODANI (1988).

Da dieser Vorgang einem Gerinnen ähnlich ist, nenne ich ihn (und damit das Programm) Coagulation. Das Programm verarbeitet in der momentanen Version die Koordinaten von 1000 Objekten und kann maximal 1000 Cluster bilden. Die Größe des .EXE-Files beträgt etwa 100.000 Bytes.

Auch hier wiederum ein komplettes Desktop, bei dem das Betrachtungsfenster aber dieses Mal die Häufigkeitsverteilung der NNED in 10 Distanzklassen als Balkendiagramm dargestellt wird.

In der Abbildung (Abb. 2) ist bereits das LEVEL-Fenster geöffnet, das es erlaubt, die Maximaldistanz (sichtbar als senkrechte Linie) zur Gruppenbildung mittels Verschieben der Maus einzustellen.

Die so gewonnene Klassifikation kann in Form zweier Datenfiles (.PNT & .CLU) abgespeichert werden, die als Eingabe für ROT3D dienen.

Beide Programme werden mit den benötigten Filenamen als Kommandozeilenparameter gestartet. Die Fehlererkennung ist mit dem entsprechenden Hilfesystem gekoppelt, sodaß z.B. bei fehlerhaften Datenfiles die Fehlerstelle angezeigt wird und somit die Fehlersuche wesentlich erleichtert wird.

© Naturwiss.-med. Ver. Innsbruck; download unter www.biologiezentrum.at

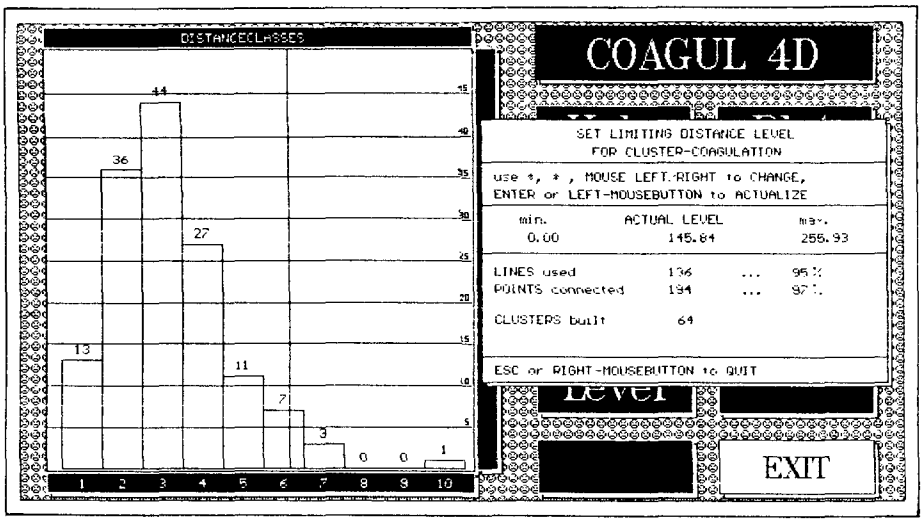

Abb. 2: Desktop von COALGUL4D (LEVEL-Fenster geöffnet).

3. Programminterna:

Die Programme sind unter Turbo Pascal(TM) 5.0 entwickelt worden (für Compilerdetails siehe SCHÄPERS, 1988). Teile der internen Datenverschiebung und der Matrixmultiplikationen (für die Rotation) sind als zeitkritische Prozeduren in Assembler-Inline Code ausgeführt (BRAUN, 1986; HARRINTON, 1988). Die Datenabspeicherung erfolgt mittels dynamischer Allokation von HeapSpeicher, das heißt, daß die Programmgröße von der verwendeten Datenmenge abhängt und nicht von vornherein bestimmt ist und damit in den meisten Fällen nur unnötig Platz verbraucht (SCHILDT, 1988). Basierend auf Turbo's Grafikroutinen wurde die komplette Fensterverwaltung selbst entwickelt. Für die Mausverwaltung wurde ein bereits beschriebenes Modul (STÜTZ, 1988) zeitlich optimiert.

Zusamenfassung: 2 Programme zur grafischen Auswertung von Ordinationsergebnissen werden vorgestellt.

ROT3D ermöglicht Rotation in Echtzeit um 3 Bildschirmachsen, Punktidentifikation durch Anklicken mit der Maus und verschiedene Verberge- und Markierungsoptionen. Die Punktkoordinaten können als Gruppen mittels einer Verbindungsliste zusammengefaßt und dargestellt werden.

COAGUL4D bildet anhand von 4 dimensionalen Punktkoordinaten eine unvollständige Minimum-Spanning-Tree-Analyse. Die Gruppengröße kann anhand der maximalen Verbindungsdistanz frei eingestellt werden.

Beide Programme benötigen einen IBM PC/AT/386(TM) (oder kompatiblen Computer) mit entweder Hercules(TM)-, EGA oder VGA-Monitor, eine Microsoft kompatible Maus, min. 512 KByte und MS-DOS(TM)2.1 oder größer. Ein 80 x 87 Coprozessor wird vorausgesetzt. Die Darstellung erfolgt auf Hercules(TM) zweifärbig (S/W), auf den anderen Grafiksystemen stehen 16 Farben zur Verfügung.

Beide Programme sind als SHAREWARE verfügbar.

Ich plane, die nächsten Versionen dieser Applikationen in eine komplette Multivariate-Analysis und Pattern-Analysis Package einzubauen, nicht nur für den PC, sondern auch für den Apple Macintosh(TM) und den Atari ST(TM). Während Sie dies lesen, sind ziemlich sicher schon Teile oder ganze Pakete verfügbar. Interessenten mögen bitte mit mir in Verbindung treten.

Copyrights: "Apple" ist ein registriertes Warenzeichen der Apple-Computer, Inc. "Atari ST" ist ein registriertes Warenzeichen der Atari Corp. "MS-DOS" ist ein registriertes Warenzeichen der Microsoft Corp.

© Naturwiss.-med. Ver. Innsbruck; download unter www.biologiezentrum.at

"GEM" ist ein registriertes Warenzeichen der Digital Research Inc. "Hercules" ist ein registriertes Warenzeichen der Hercules Corp. "IBM" ist ein registriertes Warenzeichen der IBM Corp. "Turbo Pascal" ist ein registriertes Warenzeichen der Borland International Inc.

4. Literatur:

BRAUN, U. (1986): Atari ST – 3D Grafik Programmierung. – Data Becker GmbH, Düsseldorf.

GLEICK, J. (1987): Chaos: making a niew science. – Viking Penguin Inc., New York.

HARRINGTON, S. (1988): Computergrafik – Einführung durch Programmierung. – Mc Graw Hill Inc., Hamburg, New York [et al.].

NAIMAN, A. (ed.) et al.(1988): The Macintosh Bible. – Goldstein and Blair, Berkeley.

PODANI, J. (1988): SYN-TAX III, User Manual. – Abstracta Botanica 12, Suppl. **1.**

ORLOCI, L. (1978): Multivariate Analysis in Vegetation Research (2nd ed.). – Dr. W. Junk b.v., Publ., The Hague.

SCHÄPERS, A. (Übers.) (1988): Turbo Pascal 5.0 (Benutzerhandbuch und Referenzhandbuch). – Heimsoeth software GmbH & Co. Produktions- und Vertriebs-KG, München.

SCHILDT, H. (1988): Advanced Turbo Pascal Version 4. – Borland-Osborne-Mc Graw Hill Programming Series, Berkeley.

SOUČEK, B. & M. SOUČEK (1988): Neural and massively parallel computers: the sixth generation. – John Wiley & Sons Inc., New York [et al.].

STÜTZ, B. (1988): MaNager: Turbo Pascal 4.0 unterstützt die Maus. – In: c't Magazin für Computertechnik **10:** 196 - 207.

ZOBODAT - www.zobodat.at

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: Berichte des naturwissenschaftlichen-medizinischen Verein Innsbruck

Jahr/Year: 1992

Band/Volume: 79

Autor(en)/Author(s): Hauser Michael

Artikel/Article: ROT3D & COAGULAD -2 Programme zur Auswertung von Ordinationen. 57-61