



marvin ag

AUGUR
AUGUR

Version 1.4

*Professionelle
Schrifterkennung
für den Atari ST*

Bedienungshandbuch

AUGUR
AUGUR

Marvin AG Fries-Str. 23 CH-8050 Zürich

Anmerkung

Es wurden alle erdenklichen Massnahmen getroffen, um die Richtigkeit dieser Produkt-Dokumentation zu gewährleisten. Die Marvin AG übernimmt jedoch keinerlei Garantien für den Inhalt dieses Handbuches. Im weiteren kann für keine Schäden gehaftet werden, die sich aus dem Gebrauch des Handbuches oder des Programms ergeben sollten.

Ferner behält sich die Marvin AG das Recht vor, jederzeit ohne Ankündigung Änderungen am Handbuch oder am Programm vorzunehmen.

Copyright

Die Reproduktion dieser Dokumentation, auch auszugsweise, ist ohne die schriftliche Genehmigung der Marvin AG, Zürich nicht gestattet.

Zweite, erweiterte Auflage, Januar 1989

zu AUGUR Version 1.4

Autoren des Programms AUGUR:

Michael Schenkel und Urs Müller

Autor des Handbuches:

Urs Müller

Printed in Switzerland

Inhaltsverzeichnis

Willkommen zu AUGUR.	9
Zu diesem Handbuch.	11
Beispiel-Diskette	12
1. Systemanforderungen.	13
2. Was heisst Schrifterkennung?	15
3. Über AUGUR	17
3.1. Was kann AUGUR alles lesen?	17
3.2. Wie funktioniert AUGUR?	19
4. Die Installation von AUGUR.	23
5. Vor dem Anfang.	31
5.1. Die Maus	32
5.2. Die Menuzeile	34
5.3. Warnboxen	36
5.4. Die Objektauswahl-Box	38
5.5. Das Bildfenster	42
5.6. Die Informationsfenster	44
6. Ein Beispiel.	47
6.1. Das Scannen eines Bildes	48

Erkennungsmodus	86
11. Bibliotheken	89
11.1. Eine Bibliothek laden	90
11.2. Eine Bibliothek speichern	91
11.3. Die Bibliothek löschen	92
11.4. Das Wiederverwenden gespei- cherter Bibliotheken	93
11.5. Lange Dokumente	96
12. Das Starten der Schrifterkennung	
99	
12.1. Das Segmentieren	100
12.2. Der eigentliche Erkennungs- vorgang	101
13. Der Lernmodus.	10
3	
13.1. Aufruf des Lernmodus	103
13.2. Das Übersetzen einer Bibliothek	105
13.3. Sonderzeichen	108
14. Der erkannte Text.	111
14.1. Den Text anzeigen	111
14.2. Den Text sichern	113
14.3. Den Text einer bestehenden Datei anhängen	114
14.4. Den Text über die serielle Schnittstelle senden	114
14.5. Verschiedene Textformate	119
14.6. Den Text löschen	122

Warenzeichen

AUGUR, Hawk, Hawk CP14 und Hawk 432 sind eingetragene Warenzeichen der Marvin AG.

Atari, 520ST, 520+ST, 1040ST, Mega ST, SM124 und SF354 sind Warenzeichen bzw. eingetragene Warenzeichen der Atari Corp.

GEM und GEM Desktop sind eingetragene Warenzeichen der Firma Digital Research Inc.

Ist Word Plus ist ein Warenzeichen der GST Holdings Ltd.

Macintosh ist ein Warenzeichen der Apple Computer Inc.

IBM ist ein Warenzeichen der International Business Machines Corp.

Notiz für den Anwender

Im Sinne des Fortschritts können Änderungen am Programm AUGUR oder den mitgelieferten Dateien vorgenommen werden, welche nicht im Handbuch dokumentiert sind. Solche Änderungen werden in einer Datei mit Namen "README.TXT" vermerkt. Nehmen Sie bitte diese Änderungen vor der Benutzung des Programms zur Kenntnis.

Um den Inhalt dieser Datei anzuzeigen, fahren Sie mit dem Mauszeiger auf das entsprechende Bild auf dem Bildschirm und drücken zweimal kurz hintereinander die linke Maustaste. Es erscheint eine Warnbox. Klicken Sie in das Feld "Anzeigen" und der Text erscheint auf dem Bildschirm. Mit der [Return]-Taste kann eine Zeile und mit der Leertaste eine ganze Seite vorgerückt werden. Nach der letzten Seite wird zum Desktop zurückgekehrt. Mit der Taste "Q" kann dieser Modus frühzeitig verlassen werden.

6.2.	Die Bibliothek laden	50
6.3.	Starten der Schrifterkennung . .	51
6.4.	Unbekannte Zeichen lernen . . .	52
6.5.	Den erkannten Text anzeigen . .	54
6.6.	Den Text und die Bibliothek sichern	55
7.	Das Scannen	57
7.1.	Aufstellen und Anschliessen des Scanners Hawk CP14	59
7.2.	Aufstellen und Anschliessen des Scanners Hawk 432	60
7.3.	Der Scanvorgang	64
7.4.	Das Einstellen der Helligkeit . .	66
7.5.	Das Einstellen der Auflösung . .	70
7.6.	Zur Beachtung	71
7.7.	Der IDC-Standard	72
8.	Bilder sichern und laden.	75
9.	Rahmen	77
9.1.	Rahmen einzeichnen	77
9.2.	Schrifterkennung und Rahmen . .	80
9.3.	Rahmenmarker	81
10.	Der Erkennungsmodus.	83
10.1.	Wahl des richtigen Erkennungs- modus	83
10.2.	Faustregeln für die Wahl des	

Anhang F:Mögliche Fehlerquellen 185

Anhang G: Beispiel-Dokumente. . .
. . . 191

Register 193

Willkommen zu AUGUR

Mit dem Softwarepaket AUGUR und dem Scanner Hawk CP14, Hawk 432 oder einem anderen Qualitätsscanner, für den ein IDC-Treiber existiert, besitzen Sie ein überaus vielseitiges Dokumenten-Analyse-System, mit dem Sie dank der leistungsfähigen Marvin-Technologie die verschiedensten Probleme in den Bereichen Bilderfassung, Bildmanipulation, der Schrifterkennung, sowie der Datenkompression lösen können.

Das System entstand in enger Zusammenarbeit zwischen dem Institut für Kommunikationstechnik der ETH Zürich und der Marvin AG. Es ist fähig, Dokumente mit Multifontsymbolen, Multikolonnenanordnungen, proportionalen Schriften sowie einem bunt gemischten Inhalt von Text und Grafik souverän zu verarbeiten. Seine überaus leichte Lernfähigkeit erlaubt es, das System schnell und interaktiv an unbekannte Schriftarten anzupassen.

Die gewonnenen Textfiles lassen sich mit üblichen Textverarbeitungsprogrammen, mit Desktop-Publishing-Software oder mit Datenbanksystemen weiterverarbeiten. Ferner gibt es vielfältige Möglichkeiten, reine Bilddaten zu

manipulieren, in gängige Grafikformate zu konvertieren oder zur Datenkompression zu nutzen. Grafiken oder Bilder lassen sich ebenfalls in Desktop-Publishing- Programme übernehmen.

Zu diesem Handbuch

Falls Sie noch nie mit einem Atari ST Computer gearbeitet haben, ist es von Vorteil, zuerst das Atari Bedienungshandbuch zu lesen. Danach sollten Sie das ganze vorliegende Handbuch sorgfältig studieren. Wenn Sie bereits mit dem Computersystem vertraut sind, so werden Sie leicht feststellen, welche Kapitel dieses Handbuches weggelassen werden können.

Jedes Kapitel ist so abgefasst, dass es bei Kenntnis des Atari Computers ohne die vorherigen Kapitel verstanden werden kann. Dies macht natürlich unumgänglich, dass sich das Handbuch zum Teil wiederholt.

Die Reihenfolge der einzelnen Kapitel entspricht der Abfolge der Arbeitsschritte, wie sie normalerweise beim Arbeiten mit AUGUR ausgeführt werden.

Im Anhang befinden sich verschiedene Zusammenfassungen und Tabellen. Insbesondere enthält Anhang C eine alphabetische Begriffserklärung und Anhang D eine Zusammenstellung aller möglichen Warnboxen. In Anhang F sind mögliche Fehlerquellen und **V**orschläge zu ihrer Behebung enthalten.

Beispiel-Diskette

Zusammen mit dem ROM-Modul und dem Handbuch wird auch eine Diskette mit Beispielen geliefert. Sie ist einseitig formatiert und enthält folgende Dateien:

BRIEF200.LIB	Bibliothek zu BRIEF200-.SCN
BRIEF200.SCN	Dokument "Brief", mit 200 Dpi gescannt
CONTROL.ACC	Kontrollfeld Accessory
DESKTOP.INF	Systemdatei
EMULATOR.ACC	VT52 Emulator Accessory
README.TXT	enthält aktuelle Informationen
TESTB400.LIB	Bibliothek zu TESTB400-.SCN
TESTB400.SCN	Dokument "Testbericht", mit 400 Dpi gescannt
ZÜRI200.LIB	Bibliothek zu ZÜRI200-.SCN
ZÜRI200.SCN	Dokument "Zürich", mit 200 Dpi gescannt

1. Systemanforderungen

Für den Gebrauch von AUGUR wird der Scanner Hawk CP14, Hawk 432 oder ein anderer Scanner, zu welchem ein IDC-Treiber existiert, benötigt. Weiter brauchen Sie einen Computer der Atari ST-Serie mit mindestens 1 MByte Speicher. Dieser Anforderung genügen heute folgende Modelle:

- ST 520+ (1 MByte)
- ST 1040 (1 MByte)
- Mega ST2 (2 MByte)
- Mega ST4 (4 MByte).

Computer mit 1 MByte Speicher können allerdings nur dann verwendet werden, wenn das Betriebssystem (TOS) im ROM eingebaut ist. Bei diesen Computern kann eine ganze DIN A4 Seite nur noch mit 200 Dpi gescannt werden. Bei grösseren Auflösungen kann das Dokument jedoch in mehreren Schritten mit entsprechend kleineren Ausschnitten bearbeitet werden. Ferner bleibt beim Gebrauch von AUGUR nur noch ungefähr 100 KByte Arbeitsspeicher frei, was den Gebrauch von Accessories und speicherresidenten Programmen einschränkt.

AUGUR läuft mit dem monochromen Atari-

Monitor oder mit einem professionellen Grossbildschirm. Im zweiten Fall wird von der Bildschirmfläche nur ein entsprechender linker oberer Teil gebraucht.

2. Was heisst Schrifterkennung?

Unter Schrifterkennung versteht man im weitesten Sinne die Umwandlung eines in grafischer Form gegebenen Textes in eine Vektorform. Text in grafischer Form bedeutet dabei, dass der Text nicht Buchstabe für Buchstabe gegeben ist, sondern lediglich als abgetastetes Bild, also als Matrix mit schwarzen und weissen Punkten vorliegt. Text in Vektorform heisst, dass jedem Buchstaben eine Bestimmte Zahl (Code) zugewiesen ist. Das eigentliche Aussehen der Buchstaben geht dabei verloren.

In dieser Form kann der Computer am besten mit Texten umgehen. Eine Textverarbeitung kann zum Beispiel automatische Silbentrennung und Wortkorrektur vornehmen, der Text kann nach bestimmten Kriterien sortiert oder es kann nach Wortkombinationen gesucht werden. Weiter kann dem Text eine beliebige Schriftart (Font) und Schriftgrösse verliehen werden.

AUGUR übernimmt nun die Aufgabe, Texte aus Büchern, Zeitungen, Zeitschriften, Geschäftsbriefen und vielen anderen Vorlagen in

eine für Textverarbeitungen und Desktop-Publishing-Programme verständliche Form umzuwandeln. Das heisst mit anderen Worten: AUGUR übernimmt Ihre bisherigen Abschreibearbeiten.

3. Über AUGUR

3.1. Was kann AUGUR alles lesen?

Grundsätzlich sind nur drei Anforderungen an einen Text gestellt:

1. Der Text muss gedruckt sein, das heisst gleiche Buchstaben müssen immer gleiches Aussehen haben.
2. Die Buchstaben dürfen nicht allzu oft verwachsen sein, das heisst, sie müssen normalerweise mit einer weissen Trennlinie separiert werden können.
3. Der Text sollte optisch zusammenhängen und eine einheitliche Struktur aufweisen. Einzelne Zeichen können manchmal nicht gelesen werden.

Handschriftliche Texte und Firmenlogos, sowie gewisse "Schnörkelschriften" sind also ausgeschlossen.

Weiter darf die gesamte Schrifthöhe 64 abgetastete Zeilen nicht überschreiten. Bei einer Scannerauflösung von 200 Dpi sind das 8 mm, bei einer Auflösung von 400 Dpi 4 mm. Größere Schriften müssen zuerst mit dem Pro-

gramm "SCANSOFT" (im Lieferumfang jedes Hawk Scanners enthalten) verkleinert werden oder das Dokument muss mit einer kleineren Auflösung gescannt werden.

Nach unten ist der Schriftgröße keine prinzipielle Grenze gesetzt. Bei allzu kleinen Schriften steigt die Fehlerrate immer stärker an. Als sinnvolle unterste Grenze kann eine Schrifthöhe von 16 abgetasteten Zeilen angegeben werden. Dies entspricht 2 mm bei 200 Dpi und 1 mm bei 400 Dpi. Die Grenze hängt allerdings auch von der Qualität der Vorlage (Kontrast und Schärfe) ab.

Die dritte Einschränkung wurde eingeführt, um zu verhindern, dass Schmutzflecken, Heftklammern und ähnliches den Erkennungsvorgang nicht stören oder verlangsamen.

Dies sind bereits alle Einschränkungen. Es können also auch Texte in Proportionalschrift, und in praktisch allen Schriftarten gelesen werden, Dokumente mit mehreren Schriftarten mit eingeschlossen.

3.2. Wie funktioniert AUGUR?

Zu Beginn des Erkennungsvorganges wird das Dokument auf seine Schräglage hin untersucht. Allzu schräg eingelegte Dokumente werden zurückgewiesen. Danach werden die Schriftzeilen gesucht und die darin enthaltenen Buchstaben durch weisse Linien voneinander getrennt.

Die so gefundenen Buchstaben werden anschliessend erkannt. Dazu werden die segmentierten Schriftzeichen mit Prototypen aus einer Bibliothek verglichen. Kann ein Buchstabe dort nicht gefunden werden, so wird er als neues Zeichen aufgefasst und zur Bibliothek hinzugefügt. Dabei spielt es keine Rolle, ob der Buchstabe von der selben oder von einer neuen Schriftart stammt. Deshalb können auch mehrere Schriftarten gleichzeitig erkannt werden.

Ganz am Schluss, im sogenannten Lernmodus, werden die neuen Zeichen dem Benutzer präsentiert, welcher den zugehörigen Buchstaben auf der Tastatur eingeben kann. Es ist also keine Unterbrechung des Erkennungsvorganges nötig. Treten verwachsene Zeichen auf, so können einem Zeichen auch mehrere Buchstaben zugeordnet werden (maximal vier).

Die so erweiterte Bibliothek, sowie der erkannte Text können auf Diskette gespeichert werden. Eine Bibliothek kann, abhängig von der Buchstabengrösse, der Auflösung des Scanners und der Grösse Computerspeichers maximal ungefähr 1000 Buchstaben aufnehmen. Das reicht theoretisch für ca. 10 verschiedene Schriftarten. Ist die Bibliothek aber so gross, wird der Erkennungsvorgang stark verlangsamt. Es wird empfohlen, wenn verschiedene Schriftarten nicht gemeinsam auftreten, verschiedene Bibliotheken anzulegen. Verschiedene Attribute einer Schriftart, wie Fettdruck, Kursivschrift, Schriftgrösse, werden als verschiedene Schriftarten betrachtet, müssen also einzeln gelernt werden.

Kann AUGUR mit einer vorhandenen Bibliothek nichts anfangen, das heisst es werden dort nur wenige Buchstaben gefunden, so wird der Erkennungsvorgang stark verlangsamt. Deshalb ist es im Zweifelsfalle besser eine neue Bibliothek aufzubauen (siehe nächster Abschnitt), als es mit einer vorhandenen alten Bibliothek zu versuchen.

Ist während des Erkennungsvorganges keine Bibliothek vorhanden, so baut sich AUGUR selbständig eine neue auf. Am Schluss wird jeder Buchstabe der entsprechenden Schrift

einmal gefragt werden. Somit ist AUGUR auf eine neue Schriftart trainiert. Die dazu benötigte Trainingszeit beträgt alles in allem ungefähr 5 Minuten.

4. Die Installation von AUGUR

Das Softwarepaket AUGUR wird in Cartridge-Form (Steckmodul) geliefert. Dies ist ein kleines Modul, welches Sie in Ihren Atari einstecken können und in dem das Programm hardwaremässig gespeichert ist. Da die Programm-Cartridge extern arbeitet, beansprucht AUGUR keinen Platz im Arbeitsspeicher des Computers und braucht nicht von der Diskette geladen zu werden. Das heisst: AUGUR ist immer und sofort bereit. Zudem besteht keine Gefahr von Datenverlust, wie dies bei Disketten üblich ist.

Aus dem selben Grund braucht AUGUR auch nicht installiert zu werden. Sie müssen gegebenenfalls die Cartridge in Ihrem System als Laufwerk klein "c" anmelden. Dieser Vorgang ist in Ihrem Atari- Handbuch (Kapitel 4.4.1. "Floppy anmelden") genau beschrieben. Hier folgt eine Zusammenfassung:

1. Schalten Sie Ihren Computer aus.
2. Stecken Sie die Programm-Cartridge in den Atari ein. Der Stecker befindet sich

auf der linken Seite des Computers. Die vergoldeten Kontaktstifte der Programm-Cartridge müssen dabei zum Computer zeigen. Achten Sie darauf, dass die Seite mit der Aufschrift "oben" auch wirklich nach oben zeigt. Es kann sein, dass beim ersten Mal das Einstecken etwas streng geht. Ist die Cartridge ganz eingesteckt, so verschwindet sie praktisch vollständig in der Aussparung des Cartridge-Steckers. Dabei ist sie optimal gegen mechanische Beschädigung geschützt.

3. Schieben Sie Ihre Bootdiskette mit der Aufschrift nach oben in Laufwerk A (eingebautes Laufwerk). Haben Sie noch keine Bootdiskette, so nehmen Sie die mitgelieferte Diskette.
4. Schalten Sie den Computer ein. Nach kurzer Zeit meldet sich der Atari mit dem in Abb. 4.1 dargestellten (falls Sie die mitgelieferte Diskette verwendet haben) oder einem ähnlichen Bild. Dies wird das "Desktop" (Schreibtischoberfläche) genannt. Ist auf Ihrem Bildschirm kein Laufwerkbild mit Buchstaben klein "c" sichtbar, so gehen Sie bei Punkt 5 weiter. Andernfalls gehen

Sie zu Punkt 8 und danach zu Punkt 10.



Abb. 4.1. Das Desktop

5. Klicken Sie mit der Maus auf eines der Laufwerkbilder. Mit "Klicken" ist gemeint, dass Sie den Mauszeiger (Pfeil) auf das entsprechende Objekt bewegen und dann einmal kurz die linke Maustaste drücken. Das Laufwerkbild wird daraufhin invertiert (weiss auf schwarz) dargestellt.

6. Klicken Sie jetzt auf den Menüpunkt "Floppy anmelden.." im Menu "Extras".

Bewegen Sie dazu den Mauszeiger auf das Wort "Extras" in der Menuleiste ganz oben auf dem Bildschirm. Sobald sie das Wort berühren, klappt ein Menu herunter (Abb. 4.2).



Abb. 4.2. Das Menu "Extras"

Die darin enthaltenen Einträge werden beim Überstreichen mit dem Mauszeiger jeweils invertiert dargestellt. Bewegen sie den Mauszeiger jetzt auf das Wort "Floppy anmelden.." und drücken sie einmal kurz die linke Maustaste.

7. Auf dem Bildschirm erscheint das in Abb. 4.3 dargestellte Dialogfeld. Geben Sie jetzt den Buchstaben klein "c" ("c" steht für Cartridge) über die

Tastatur ein. Um den Bildtitel zu verändern, zeigen Sie mit dem Mauszeiger auf den derzeit vorgegebenen Namen und drücken Sie einmal kurz die linke Maustaste. Um den Namen zu löschen, drücken Sie die Taste [Backspace] oder [Esc] auf der Tastatur. Jetzt können Sie einen beliebigen Namen eintippen, zum Beispiel "ROM-MOODUL" Dieser Name hat auf die Funktionsweise keinen Einfluss, er dient Ihnen lediglich als Gedankenstütze. Um Ihre Eingabe zu bestätigen klicken Sie in das Feld "Anmelden".



Abb.4.3. Das Dialogfeld "Floppy anmelden"

8. Auf dem Bildschirm ist jetzt ein Laufwerk mit dem Buchstaben klein "c" zu sehen. Zeigen Sie mit der Maus auf dieses Laufwerk und drücken Sie zwei

Mal sehr kurz hintereinander die linke Maustaste (Doppelklick). Daraufhin öffnet sich ein Fenster, in dem ein Bild mit dem Namen "AUGUR" sichtbar wird (Abb. 4.1).

9. Zeigen Sie mit dem Mauszeiger wieder auf das Menu "Extras" und klicken Sie diesmal auf den Eintrag "Arbeit sichern". Eventuell folgt dann eine Sicherheitsabfrage des Systems, die Sie mit Klicken in das "Ja"-Feld bestätigen. Die soeben vorgenommenen Modifikationen werden jetzt auf Diskette gespeichert. Wenn Sie beim nächsten Einschalten die selbe Diskette in Ihren Computer einlegen, erscheint das Laufwerkbild klein "c" sofort. Falls Ihre Diskette schreibgeschützt war, fordert Sie das System auf, den Schreibschutz zu entfernen (kleiner Schieber oben rechts auf Ihrer Diskette so schieben, dass er das Loch verdeckt).
 10. Durch Doppelklick in das "AUGUR"-Bild kann AUGUR jetzt gestartet werden (Bewegen Sie den Mauszeiger auf das "AUGUR"-Bild und drücken Sie die linke Maustaste zweimal in rascher Abfolge).
-

Hinweis: Besitzt Ihr Atari nur 1 MByte Arbeitsspeicher (Atari 1040ST oder 520+ST), so verbleiben neben AUGUR noch etwa 100 KByte Speicher frei. Haben Sie in Ihrem System speicherplatzraubende residente Programme und Accessories installiert, so kann es vorkommen, dass für AUGUR zuwenig Platz bleibt. Beim Versuch AUGUR zu starten erscheint eine Warnbox mit einer entsprechenden Meldung. In diesem Fall müssen Sie einige dieser residenten Programme oder Accessories entfernen.

5. Vor dem Anfang

Am Schluss des letzten Kapitels haben Sie AUGUR gestartet. Abb. 5.1. zeigt den Bildschirm, wie er sich danach präsentiert.

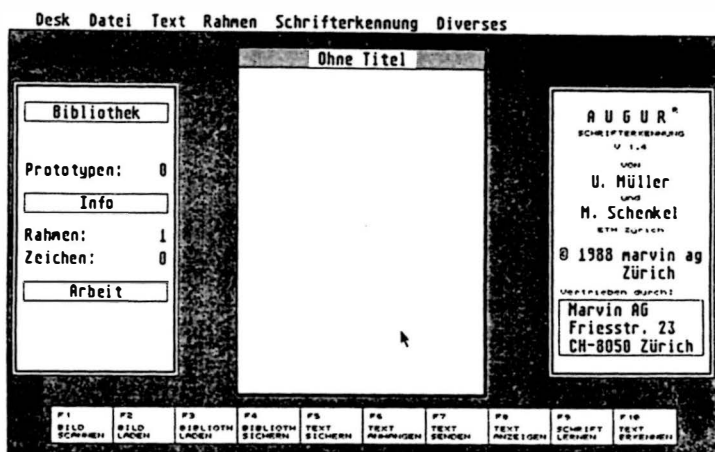


Abb. 5.1. Der Bildschirmaufbau von AUGUR

AUGUR arbeitet unter GEM (Graphics Environment Manager). Dies ist ein Grafiksystem, welches Ihnen die Anwendung von Programmen erleichtert. Einerseits haben alle Programme welche unter GEM laufen einen

ähnlichen Aufbau und Sie brauchen diesen nur einmal zu erlernen. Zum anderen entfällt das lästige Auswendiglernen von Befehlssequenzen.

Im folgenden werden die wenigen Regeln erklärt, welche Sie für das Arbeiten mit GEM und für AUGUR benötigen. Viele dieser Regeln finden Sie auch in Ihrem Atari Bedienungshandbuch, dort zum Teil detaillierter. An geeigneter Stelle wird jeweils auf dieses Handbuch verwiesen werden.

5.1. Die Maus

Grundlegend für das Arbeiten mit GEM-Programmen ist die Maus. Sie hat auf ihrer Unterseite ein Kugel, welche sich beim Schieben der Maus auf einer ebenen Fläche dreht. Anhand dieser Drehung stellt der Computer die Bewegungsrichtung und -geschwindigkeit fest. Normalerweise befindet sich auf dem Bildschirm ein Pfeil, welcher die Bewegungen, die Sie mit der Maus ausführen, nachvollzieht (Mauszeiger). Achten Sie darauf, dass Sie die Maus nicht verdreht oder schief halten, sonst stimmt die Bewegungsrichtung des Pfeils auf dem Bildschirm nicht mit jener der Maus überein.

Kapitel 5: Vor dem Anfang

Geraten Sie beim Bewegen der Maus an ein Hindernis, so können Sie durch Anheben der Maus (die Kugel dreht dann nicht) neu ansetzen.

Die Maus besitzt vorne zwei Tasten. Bewegen Sie den Pfeil auf dem Bildschirm auf gewisse Objekte und drücken Sie die linke Maustaste, so wird eine entsprechende Aktion ausgeführt. Die rechte Maustaste wird sehr selten gebraucht.

Wenn der Computer am Arbeiten ist, erscheint an Stelle des Pfeils eine (fleissige) Biene.

Über die Bedienung der Maus hat sich folgendes Vokabular eingebürgert:

klicken: Drücken (und Loslassen) der linken Maustaste.

Doppelklick: In schneller Abfolge zweimal die linke Maustaste drücken. Die Geschwindigkeit, in welcher dieser Doppelklick zu erfolgen hat lässt sich einstellen. Nähere Informationen entnehmen Sie bitte Ihrem Atari-Handbuch (Kapitel 4.1.3. "Kontrollfeld").

auswählen: Den Mauszeiger auf das entsprechende Objekt bewegen und die linke

Maustaste drücken.

5.2. Die Menuzeile

Ganz zuoberst auf dem Bildschirm befindet sich die sogenannte Menuzeile. Sie enthält bei AUGUR sechs verschiedene Einträge. Es handelt sich um sogenannte Dropdown-Menüs (herunterfallende Menüs). Fahren Sie mit der Maus auf einen dieser Einträge so fällt ein Untermenü mit weiteren Einträgen herunter. Bewegen Sie den Mauszeiger abwärts, so werden die Menüeinträge unter dem Mauszeiger invertiert (weiss auf schwarz) dargestellt. Drücken Sie die linke Maustaste, so wird der gerade überstrichene Befehl ausgeführt. Wollen Sie keinen Befehl ausführen, so bringen Sie das Menü wieder zum verschwinden, indem Sie die Maus ganz aus dem Menü herausfahren und die linke Maustaste drücken.

Als Beispiel klicken Sie auf den Menüeintrag "AUGUR..." im Menü "Desk". Es erscheint die sogenannte Infobox mit Angaben über das Programm und den Distributor (Abb. 5.2).

Sie bringen diese Box durch Klicken in das "OK" Feld wieder zum Verschwinden.



Abb. 5.2. Die Infobox

Einige häufig gebrauchte Befehle können Sie auch mit den Funktionstasten auslösen. Die jeweiligen Funktionstasten sind in den entsprechenden Menueinträgen vermerkt. Zusätzlich sind Symbole der zehn Funktionstasten auf dem unteren Bildschirmrand abgebildet. Eine dritte Alternative, diese Befehle auszulösen, besteht darin, mit der Maus in das Grafiksymbol der entsprechenden Funktionstaste zu klicken.

Im folgenden wird jeweils nur noch auf die erste Möglichkeit hingewiesen.

Während der Ausführung eines Befehles wird der entsprechende Menutitel invertiert (weiss auf schwarz) dargestellt. Handelt es sich um eine Funktion, welche auch über eine der Funktionstasten aufgerufen werden kann, so wird auch das entsprechende Grafiksymbol am unteren Bildschirmrand invertiert dargestellt.

5.3. Warnboxen

Warnboxen sind umrahmte Rechtecke, welche manchmal in der Bildschirmmitte auftauchen. Sie werden von AUGUR benützt, um Sie vor einem Schritt mit Konsequenzen gegebenenfalls zu warnen oder um Sie über einen bestimmten Sachverhalt zu informieren.

In der Warnbox befinden sich ein oder mehrere Antwortknöpfe. Wenn Sie mit der Maus in einen dieser Knöpfe klicken, wird die Warnbox entfernt und AUGUR führt das Programm in entsprechender Weise fort. Den fett umrahmten Knopf können Sie auch durch Drücken von [Return] oder [Enter] anwählen.

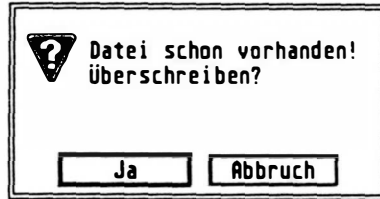


Abb. 5.3. Beispiel einer Warnbox

Ein Beispiel ist in Abb. 5.3. dargestellt. Diese Warnbox erscheint immer dann, wenn versucht wird eine Datei unter einem bereits bestehenden Namen abzuspeichern. Drücken Sie "JA", so wird die alte Datei durch die neue ersetzt, also gelöscht. Drücken Sie "Abbruch", wird der Speichervorgang abgebrochen.

Bei normalem Programmablauf treten Warnboxen nur selten auf. Nicht alle Warnboxen, welche bei einer bestimmten Situation auftreten können, sind deshalb in diesem Handbuch erwähnt. Der Text würde damit nur unlesbar gemacht. Eine vollständige Auflistung aller Warnboxen und deren Bedeutung finden Sie im Anhang D ("Warnboxen").

5.4. Die Objektauswahl-Box

Die Objektauswahl-Box ist Teil des Atari Betriebssystemes. Sie erscheint immer vor dem Laden und Speichern von Dateien und dient zum Auswählen der gewünschten Dateinamen (auch in anderen Programmen). Im Handbuch des Mega ST Computers ist die Objektauswahl-Box im Kapitel 5.2.1 "Benutzen der Objektauswahl-Box" detailliert beschrieben. Im Handbuch zu den Computern 1040ST und 520+ST fehlt dieses Kapitel leider. Im folgenden sind die für AUGUR notwendigen Manipulationsmöglichkeiten beschrieben. Schieben Sie die mitgelieferte Beispiel-Diskette in das eingebaute Laufwerk (Laufwerk A) und klicken Sie dann auf den Menüeintrag "Bild laden..." im Menu "Datei". Es erscheint die Objektauswahl-Box, wie in Abb. 5.4 dargestellt.

In einem weiteren Fenster innerhalb der Objektauswahl-Box wird das Inhaltsverzeichnis der aktuellen Diskette angezeigt. In diesem Inhaltsverzeichnis sind nur jene Dateien sichtbar, welche den selben Extender (Erweiterung des Dateinamens) haben wie rechts in der Indexzeile angegeben. Dies ist der besseren Übersicht sehr dienlich. Wenn Sie zum Beispiel eine Bibliothek laden wollen, werden im

Inhaltsverzeichnis auch nur Bibliotheken angezeigt.

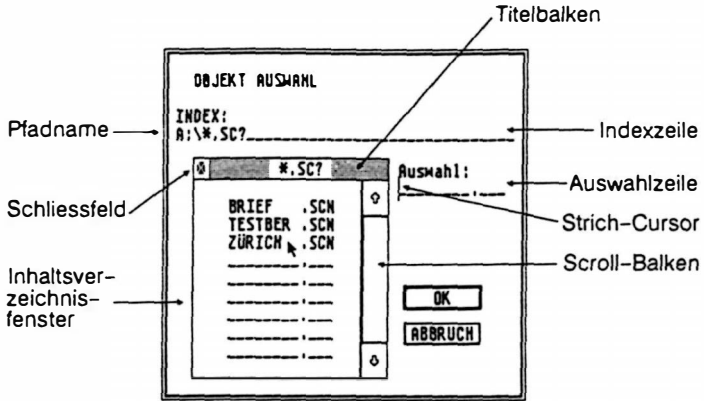


Abb. 5.4. Die Objektauswahl-Box

Sie haben drei Möglichkeiten eine Datei auszuwählen:

Klicken Sie mit der Maus auf den gewünschten Dateinamen. Er wird daraufhin invertiert dargestellt und in die Auswahlzeile übertragen. Klicken Sie jetzt auf das "OK"-Feld oder drücken Sie [Return] oder [Enter].

Dasselbe Resultat können Sie auch mit einem

Doppelklick auf den gewünschten Dateinamen erzielen.

Die dritte Möglichkeit besteht darin, den gewünschten Namen in der Auswahlzeile über die Tastatur einzugeben und dann auf "OK" zu klicken. Dies ist beim Anlegen einer neuen Datei (der Name erscheint also noch nicht im Inhaltsverzeichnis) der einzige Weg.

Beinhaltet ein Inhaltsverzeichnis mehr Dateien, als zugleich in der Anzeige Platz finden, so können diese durch Verschieben des Fensterausschnittes sichtbar gemacht werden. Nehmen Sie dazu den weissen Scroll-Balken am rechten Rand des Inhaltsverzeichnisses (linke Maustaste gedrückt halten) und verschieben Sie ihn in die gewünschte Richtung. Eine zweite Möglichkeit besteht darin, in einen der beiden Pfeile ober- und unterhalb des Scrollbalkens zu klicken. Das Inhaltsverzeichnis wird dann zeilenweise in die entsprechende Richtung verschoben.

Ordner werden im Inhaltsverzeichnis mit einem rechteckigen Zeichen links vom Dateinamen gekennzeichnet (Ordner sind im Atari-Handbuch, Kapitel 4.2 "Das Menü 'Datei'" beschrieben). Klicken Sie auf einen solchen Ordnernamen, so werden jene Dateien ange-

zeigt, welche sich in dem entsprechenden Ordner befinden. Das Unterverzeichnis wird durch Klicken in das Schliessfeld (links vom Titelbalken des Inhaltsverzeichnis) wieder verlassen.

Wenn Sie die Diskette wechseln, klicken Sie in den Titelbalken des Inhaltsverzeichnis um den Inhalt der neuen Diskette anzuzeigen.

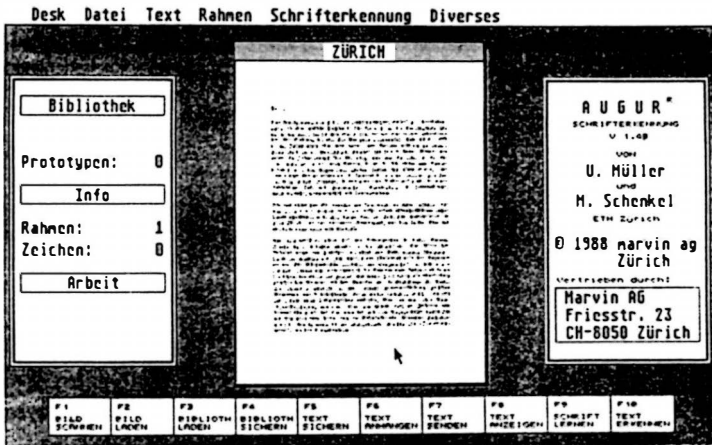


Abb. 5.5. Das Bild "ZÜRICH" wurde geladen

Oft wird von AUGUR in der Auswahlzeile bereits ein Name vorgeschlagen. Sind Sie da-

mit einverstanden, so müssen Sie nur noch in das "OK"-Feld klicken oder [Return] drücken.

Laden Sie für das nächste Beispiel das Bild "ZÜRI200.SCN" (Abb. 5.5).

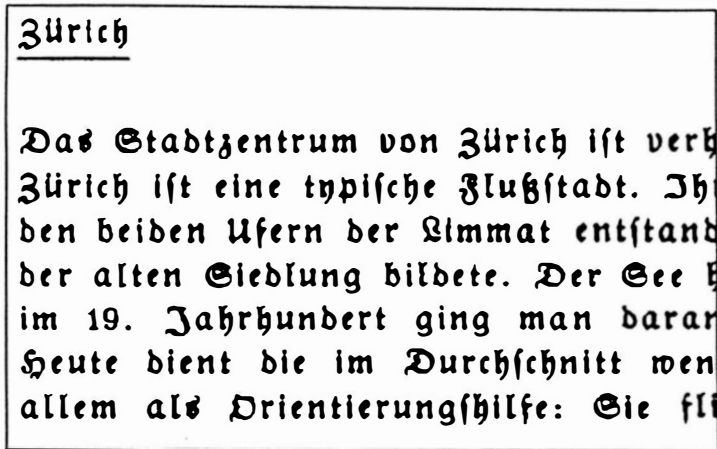
5.5. Das Bildfenster

In der Bildschirmmitte befindet sich ein Fenster, welches eine Verkleinerung des aktuellen Bildes enthält. Im Balken darüber steht der Name des Bildes (Abb. 5.5). Trägt das Bild noch keinen Namen (weil es noch nie abgespeichert wurde), steht dort "Ohne Titel" (Abb. 5.1).

Sie können einzelne Bildausschnitte auch vergrößert betrachten. Fahren Sie dazu mit der Maus auf die entsprechende Stelle im verkleinerten Bild und klicken Sie mit der linken Maustaste. Der entsprechende Ausschnitt wird vergrößert auf dem ganzen Bildschirm angezeigt (Abb. 5.6).

Bewegen Sie jetzt die Maus, so bewegt sich der sichtbare Ausschnitt in entsprechender Richtung (das Bild "scrollt"). Sie können damit das ganze Bild betrachten. Dieser Modus

wird "Scroll-Modus" genannt. Er kann durch Drücken einer Taste (Tastatur oder Maus) verlassen werden. Der Mauszeiger steht im verkleinerten Bild dann an der zuletzt betrachteten Stelle.



*Abb. 5.6. Vergrößerter Bildausschnitt
(Scroll-Modus)*

Einige der Befehle, welche auch über die Funktionstasten aufrufbar sind (Insbesondere das Starten der Schrifterkennung), können ebenfalls vom Scrollmodus aus ausgeführt werden. Drücken Sie dazu im Scrollmodus einfach die entsprechende Funktionstaste. Der

Scrollmodus wird verlassen und der entsprechende Befehl wird ausgeführt.

5.6. Die Informationsfenster

Links und rechts auf dem Bildschirm befinden sich zwei weitere Fenster. Im rechten sind Angaben über die Autoren von AUGUR und im unteren Teil die Adresse Ihres Händlers eingetragen.

Das linke Fenster enthält die drei Bereiche "Bibliothek", "Info" und "Arbeit". Deren Bedeutung wird an geeigneter Stelle beschrieben werden. Trotzdem seien sie hier kurz erklärt:

Unter dem Titel "Bibliothek" steht der Name der jeweils geladenen Bibliothek. Ist diese erweitert und noch nicht auf Diskette zurückspeichert, so steht rechts davon das Wort "neu". Rechts vom Eintrag "Prototypen:" steht wieviele verschiedene Zeichen sich in dieser Bibliothek befinden.

Der Titel "Info" enthält zwei Einträge: "Rahmen:" zeigt an wieviele Rahmen insgesamt eingezeichnet sind. "Zeichen:" besagt, wieviele Buchstaben sich im Textspeicher

befinden, also erkannt worden sind.

Der Titel "Arbeit" schliesslich kann verschiedene Einträge enthalten, je nach Arbeit, welche gerade ausgeführt wird.

6. Ein Beispiel

In diesem Kapitel soll anhand eines Beispiels das Arbeiten mit AUGUR erläutert werden. Die 6 Unterkapitel entsprechen dabei den einzelnen Arbeitsschritten, wie sie üblicherweise beim Erkennen von Dokumenten ausgeführt werden. Eine detaillierte Beschreibung aller Funktionen folgt in den nächsten Kapiteln.

Falls Sie es nicht bereits getan haben, schließen Sie jetzt Ihren Scanner an den Atari an. Beachten Sie dazu Kapitel 7.1. oder 7.2. Wenn Sie einen anderen als den Hawk CP14 oder Hawk 432 Scanner gebrauchen möchten, so benötigen Sie einen entsprechenden IDC-Treiber. Lesen Sie in diesem Fall bitte zuerst die dem Treiber beigelegten Blätter zur richtigen Handhabung des Scanners durch.

In den letzten Kapiteln wurde gezeigt, wie AUGUR gestartet und wie ein Bild geladen wird. Haben Sie AUGUR noch nicht gestartet, so tun Sie das jetzt.

6.1. Das Scannen eines Bildes

Nehmen Sie aus dem Anhang G dieses Handbuchs das Beispieldokument 1 (Geschäftsbrief) und legen Sie es unter den Scanner.

Beim Hawk Scanner müssen Sie den Schlitten von Hand ganz nach rechts schieben und den Helligkeitsregler auf Position 3 bis 4 stellen.

Der Hawk 432 beherrscht mehrere Auflösungen. Sie sollten ihn für dieses Beispiel auf 200 Dpi schalten. Wählen Sie zu dazu den Punkt "Scanner..." im Menu "Diverses" an. Es erscheint ein Dialogfeld in dem Sie Scannertyp, Auflösung und Helligkeit wählen können (Abb. 6.1). Klicken Sie zuerst auf das Feld "200" unter dem Titel "Auflösung:" und danach auf das Feld "OK". Die Dialogbox verschwindet wieder.

Klicken Sie jetzt auf den Menüpunkt "Bild scannen" im Menu "Datei". Der Hawk 432 Scanner startete automatisch. Beim Hawk CP14 muss noch der Startknopf gedrückt werden.

In ungefähr 10 Sekunden wird das Dokument abgetastet und in den Computer gelesen. Danach erscheint die gewohnte Verkleinerung des

Dokumentes im Fenster der Bildschirmmitte. Betrachten Sie das Dokument auch im Scroll-Modus (mit der Maus in das verkleinerte Bild klicken). Falls das Dokument schief eingelegt war, auf dem Kopf steht oder Teile davon abgeschnitten sind, wiederholen Sie den Scanvorgang.

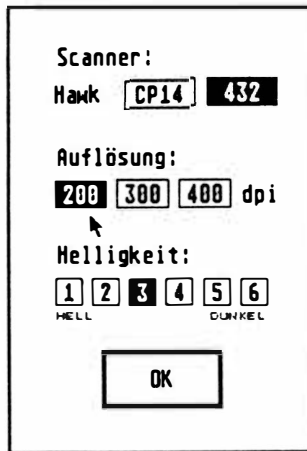


Abb. 6.1. Wenn Sie einen Hawk 432 Scanner besitzen, so können Sie mit diesem Dialog die Auflösung und Helligkeit wählen.

6.2. Die Bibliothek laden

In einer Bibliothek sind Prototypen zu jedem Buchstaben der entsprechenden Schrift enthalten. Laden Sie die richtige Bibliothek mit dem Befehl "Bibliothek laden" unter dem Menu "Datei". Es erscheint die Objektauswahl-Box. In unserem Fall benötigen wir die Bibliothek mit Namen "BRIEF200.LIB". Klicken Sie mit der linken Maustaste auf diesen Namen und anschliessend auf das "OK" Feld (Abb. 6.2). Bei der Verwendung eines anderen Scanners als dem Hawk CP14 oder 432 ist es besser, keine der obigen Bibliotheken zu laden. Die Bibliothek wird dann vollständig neu aufgebaut.

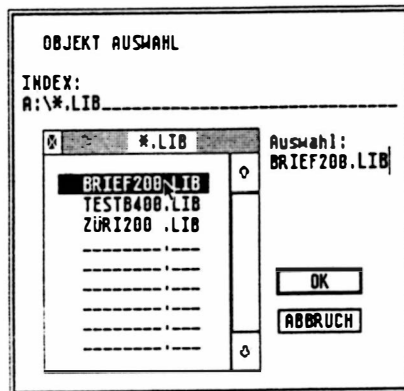


Abb. 6.2. Auswählen der richtigen Bibliothek

Ist die Bibliothek geladen, so erscheint ihr Name links oben auf dem Bildschirm unter dem Titel "Bibliothek".

6.3. Starten der Schrifterkennung

Zum Starten der Schrifterkennung klicken Sie auf den Befehl "Text erkennen" unter dem Titel "Schrifterkennung".

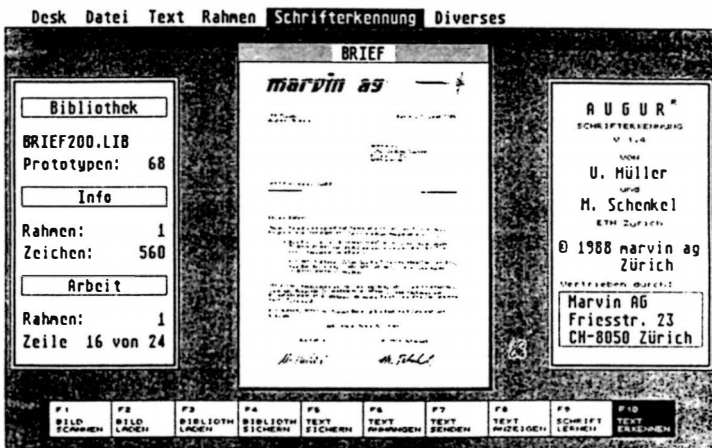


Abb. 6.3. Der Erkennungsvorgang

Unter dem Titel "Arbeit" erscheint zuerst die

Meldung "Linien entfernen". AUGUR entfernt alle horizontalen und vertikalen Striche. Vor allem die Unterstreichung des Wortes "Zürich" in der Adresse könnte sich eventuell störend auf den Erkennungsvorgang auswirken. Der Arbeitsschritt "Linien entfernen" ist jedoch nicht zwingend, in der geladenen Bibliothek war jedoch vermerkt, dass AUGUR diesen Schritt ausführen soll. Mehr darüber erfahren Sie in Kapitel 15.6. "Linien entfernen".

Ist das Linien Entfernen beendet, so folgt die Meldung "Segmentieren". Damit wird angezeigt, dass AUGUR im Moment den Ort und die Grösse der vorhandenen Schriftzeilen sucht. Danach beginnt der eigentliche Lesevorgang. Die Anzeige "Segmentieren" erlischt, anstelle wird angezeigt, wieviele Schriftzeilen gefunden wurden, und welche im Moment gerade bearbeitet wird. Unter dem Titel "Info" wird angezeigt, wieviele Buchstaben bereits gelesen worden sind. Der Titel "Prototypen:" (unter "Bibliothek") zeigt an, wieviele Zeichen sich in die Bibliothek befinden (Abb. 6.3).

6.4. Unbekannte Zeichen lernen

Falls im Text unbekannte Zeichen gefunden

6.5. Den erkannten Text anzeigen

Falls keine neuen Zeichen gefunden wurden, oder nach Beenden des Lernmodus, wird der gelesene Text auf dem Bildschirm angezeigt (Abb. 6.5).

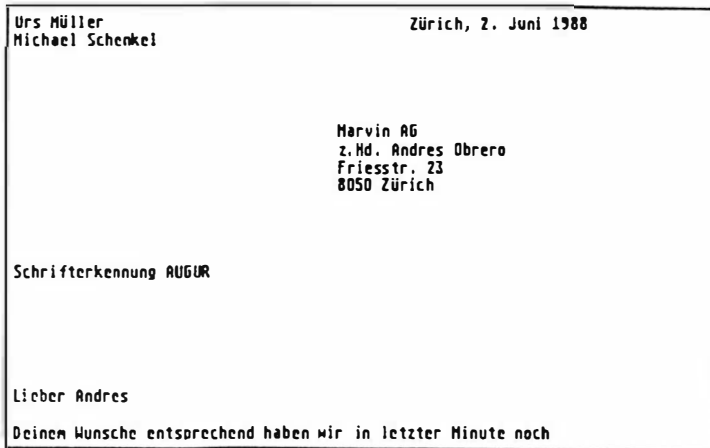


Abb. 6.5. Der erkannte Text

Vergleichen Sie den gelesenen Text mit dem Original. Auf einen Tastendruck (Tastatur oder Maus) wird jeweils die nächste Bildschirmseite angezeigt. Mit einem Tastendruck nach der letzten Seite oder mit [Esc] wird dieser Modus wieder verlassen.

6.6. Den Text und die Bibliothek sichern

Mit dem Menüpunkt "sichern..." im Menü "Text", kann der erkannte Text auf Diskette gespeichert werden, um ihn anschliessend mit einem Textverarbeitungsprogramm weiterverarbeiten zu können. Mit dem Menüpunkt "Bibliothek sichern" im Menü "Datei", kann die erweiterte Bibliothek wieder zurückgeschrieben werden, damit die neu gelernten Buchstaben bei Beendigung des Programmes nicht verloren gehen. Falls Sie es wünschen, kann auch das Originalbild abgespeichert werden ("Bild sichern") im Menü "Datei".

7. Das Scannen

Unter Scannen (auf Deutsch: Abtasten) versteht man das Umwandeln eines Bildes in eine für den Computer verständliche Form. Im vorliegenden Fall wird das Original in eine feine Matrix von Punkten (8 bis 16 Punkte pro Millimeter, je nach Auflösung des verwendeten Scanners) unterteilt, wobei jeder Punkt die Farbe Schwarz oder Weiss haben kann.

Wichtig zu wissen ist, dass die richtige Einstellung des Scanners grossen Einfluss auf die Qualität des Erkennungsprozesses haben kann.

AUGUR ist speziell für die Scanner Hawk CP14 und Hawk 432 der Firma Marvin AG konzipiert worden.

Die besten Resultate werden mit dem Hawk 432 erzielt. Er beherrscht die Auflösungen 200, 300 und 400 Dpi (entspricht 8, 12 und 16 Punkten/mm) und lässt sich vollständig über die Software steuern.

Der Hawk CP14 kennt nur die Auflösung 200 Dpi, besitzt dafür einen eingebauten Thermo drucker und kann gegebenenfalls auch als Fotokopierer verwendet werden. Die Helligkeit

muss auf dem Bedienungsfeld eingestellt werden (blauer Knopf). Von dort wird er ebenfalls gestartete (roter Knopf).

AUGUR kann jedoch auch andere Scanner bedienen. Zu diesem Zweck wurde der **IDC-Standard** geschaffen (siehe Kapitel 7.6).

AUGUR detektiert beim Starten welcher Scanner (Hawk CP14, 432 oder ein IDC-Treiber) angeschlossen ist und schaltet sich automatisch auf diesen. Mit dem Befehl "Bild Scannen" im Menu "Datei" (siehe weiter unten) wird danach automatisch der richtige Scanner angesprochen.

Sind Sie im Besitz eines Hawk CP14 so lesen Sie bitte das Kapitel 7.1 und gehen danach zu Kapitel 7.3. Besitzen Sie einen Hawk 432, so lesen Sie bitte bei Kapitel 7.2 weiter. Besitzen Sie einen anderen Scanner und möchten diesen über IDC mit AUGUR verwenden, so studieren Sie bitte aufmerksam die dem IDC-Treiber beigelegten Blätter und gehen dann zu Kapitel 7.3.

7.1. Aufstellen und Anschliessen des Scanners Hawk CP14

Für eine genaue Beschreibung über das Auspacken, Aufstellen, die Inbetriebnahme und die Pflege des Scanners lesen Sie bitte im entsprechenden Scanner-Handbuch nach.

Stellen Sie den Scanner auf eine horizontale Fläche. Achten Sie darauf, dass der Auflage-schlitten (Originalauflage) über genügend Bewegungsfreiheit verfügt. Er muss von Hand bis zum Anschlag nach links und auch nach rechts bewegt werden können, ohne dabei einen Gegenstand zu berühren.

Verbinden Sie den Scanner mit dem Atari über ein Centronics-Kabel (Drucker-Kabel, im Lieferumfang des Scanners enthalten). Beide Geräte sollten dabei ausgeschaltet sein. Die entsprechenden Stecker befinden sich auf der Rückseite von Scanner und Computer. Der richtige Steckplatz am Atari ist mit "Printer" beschriftet. Schliessen Sie auch das Netzkabel an den Scanner an. Schalten Sie danach zuerst den Scanner (grünes Lämpchen leuchtet) und dann den Computer ein.

Hinweis: Damit beim Abtasten (Schlitten fährt von rechts nach links) keine Ver-

zerrungen auftreten, ist der Scanner mit einem starken Motor ausgerüstet. Vor allem Gegenstände, welche links vom Scanner abgestellt sind (zum Beispiel Gläser) laufen Gefahr, beim Abtasten vom Schreibtisch geworfen zu werden.

Zum Abtasten übergrosser oder dicker Originale besteht die Möglichkeit, die Abdeckung des Scanners zu entfernen. Greifen Sie dazu (bei geschlossenem Deckel) mit beiden Daumen in die Einbuchtung des Gelenkes am Hinterteil der Abdeckung. Drücken Sie fest nach hinten, wobei Sie mit den restlichen Fingern den Scanner festhalten. Die Abdeckung fällt heraus. Achten Sie beim Wiedereinsetzen auf die Aussparungen in der Auflagefläche.

Richten Sie die Dokumente immer nach der vorderen und der rechten Kante der Auflagefläche aus (an der rechten vorderen Ecke befindet sich ein Pfeil).

7.2. Aufstellen und Anschliessen des Scanners Hawk 432

Für eine genaue Beschreibung über das Aus-

packen, Aufstellen, die Inbetriebnahme und die Pflege des Scanners lesen Sie bitte im entsprechenden Scanner-Handbuch nach.

Stellen Sie den Scanner auf eine horizontale Fläche und achten Sie darauf, dass die Transportsicherung (siehe Handbuch des Hawk 432 Scanners) entfernt worden ist.

Schliessen Sie den Scanner mit dem Netzkabel an einer Steckdose an. Achten Sie darauf, dass der Scanner noch ausgeschaltet ist. Der Netzschalter des Hawk 432 befindet sich auf der rechten Seite, ganz vorne.

Schalten Sie den Atari aus. Stecken Sie die Interfacebox auf der Rückseite des Computers ein. Der richtige Steckplatz ist mit "HardDisk" beschriftet (die Interfacebox weist insgesamt vier Stecker auf, nur einer davon passt in diesen Steckplatz).

Die Interfacebox muss noch mit Strom versorgt werden. Für die Computer der Mega ST Serie wird dazu ein spezielles Spiralkabel für die Tastatur geliefert, welches an einem Ende ein dünnes Kabel mit Anschlussstecker für die Interfacebox enthält. Ersetzen Sie das alte Spiralkabel durch dieses, wobei das Ende mit dem dünnen Kabel im Computer eingesteckt

sein sollte. Den Stecker am dünnen Kabel verbinden Sie jetzt mit der Interfacebox (nur ein Steckplatz passt).

Für die Computer 1040 ST und 520+ST ist diese elegante Variante leider nicht möglich, da die Tastatur ja im Computer eingebaut ist. Für diese Fälle wird ein Stromkabel geliefert, welches am Joystick-Port des Computers (beim 1040 unten, beim 520+ auf der rechten Seite) eingesteckt werden muss.

Bei einigen Modellen des Hawk 432 wird der Strom direkt vom Scanner geliefert. In diesem Fall wird gar kein zusätzliches Stromkabel benötigt (und auch nicht geliefert).

Ist die Stromverbindung für die Interfacebox hergestellt, so verbinden Sie den Scanner und den Atari über das mitgelieferte Anschlusskabel. Das eine Ende wird an der Interfacebox angeschlossen (nur ein Steckplatz passt), das andere Ende an der Rückwand des Hawk 432.

An der Interfacebox bleibt jetzt noch ein Steckplatz frei, an dem weitere Geräte, wie Harddisk oder Laserdrucker angeschlossen werden können. Ist dies der Fall und haben Sie einen Scanner, der den Strom für die Interfacebox selbst liefert, so muss der Scanner,

beim Arbeiten mit dem Computer, immer eingeschaltet sein, da die Interfacebox sonst keinen Strom erhält und die Daten demzufolge nicht richtig weiterleiten kann. Sie können aber auch die Anschlussgeräte, wie früher, direkt in den Atari einstecken und die Interfacebox an das letzte Gerät dieser Kette anschliessen. Da in diesem Fall kein weiteres Gerät folgt, muss die Interfacebox nicht dauernd mit Strom versorgt werden und auch Scanner, welche den Strom für die Interfacebox liefern müssen nicht dauernd eingeschaltet bleiben.

Sind die Anschlüsse alle gemacht, so schalten Sie zuerst den Scanner (grünes Lämpchen leuchtet) und dann den Computer ein.

Zum Abtasten übergrosser oder dicker Originale besteht die Möglichkeit, die Abdeckung des Scanners zu entfernen. Klappen Sie dazu die Abdeckung nach oben. Fassen Sie diese auf beiden Seiten möglichst weit unten und schieben sie nach rechts. Nachdem die Abdeckung aus ihren Scharnieren gesprungen ist, heben Sie diese vorsichtig nach oben weg. Achten Sie beim Wiedereinsetzen auf die Aussparungen in der Auflagefläche.

Richten Sie die Dokumente immer nach der

hinteren und der rechten Kante der Auflagefläche aus.

7.3. Der Scanvorgang

Falls dies nicht bereits geschehen ist, starten Sie jetzt das Programm AUGUR. Beim Scannen können Sie zwischen Hoch- und Querformat wählen. Dies geschieht mit den Menüpunkten "**Hochformat**" und "**Querformat**" im Menu "Diverses". Das gewählte Format wird mit einem Haken versehen und das Fenster mit dem verkleinerten Bild passt sich dem jeweiligen Format an.

Das Hochformat dient für allgemeine Dokumente, Zeitschriften und Zeitungen. Das Querformat ist vor allem für Bücher gedacht. Viele Bücher sind klein genug, damit zwei Buchseiten gemeinsam auf dem Scanner Platz finden, wenn das Buch quer darauf gelegt wird.

Der Scanner Hawk 432 liefert nicht direkt das Querformat. Deshalb muss das Bild im Computerspeicher in diesem Modus zuerst gedreht werden. Während dieser Zeit erscheint unter dem Titel Arbeit die Meldung "Bild rotieren". Bei der grossen Datenmenge dauert dieser

Vorgang ungefähr 15, 45 oder 95 Sekunden für eine gescannten DIN A4 Seite von 200, 300 oder 400 Dpi. Zusätzlich erschwerend ist die Tatsache, dass kein zweiter Bildspeicherplatz reserviert werden kann und das Bild deshalb "auf sich selbst" gedreht werden muss. Das Bild rotieren kann durch einen Tastendruck der Tastatur abgebrochen werden.

Ausgelöst wird der Scanvorgang mit dem Menüpunkt "Bild scannen" im Menü "Datei". Der Mauszeiger wechselt in eine Biene.

Der Hawk 432 startet jetzt automatisch. Beim Hawk CP14 müssen Sie den Auflageschlitten zuerst von Hand bis zum rechten Anschlag schieben und danach die Starttaste (Rosa Knopf am Scanner) drücken. Wenn Sie sich vor dem Drücken der Starttaste des Hawk CP14 umentschieden haben, so können Sie den Scanvorgang durch Drücken einer Taste der Tastatur abbrechen. Der Scanvorgang dauert beim Hawk CP14 ungefähr 10 Sekunden, beim Hawk 432 dauert er 7 bis 17 Sekunden, je nach gewählter Auflösung (Siehe Kapitel 7.5).

Beim Hawk CP14 müssen Sie für kleinere Vorlagen den Schlitten nicht ganz nach rechts schieben. Wichtig ist nur, dass sich der untere Rand des Dokumentes rechts von der Fluores-

zenlampe befindet. Es befindet sich eine entsprechende Markierung ("Start Position") auf der Frontseite des Scanners. Auf der Vorderseite des Auflageschlittens ist ein beweglicher Cursor angebracht. Schieben Sie diesen auf die Markierung, so ist ein anschliessend korrektes Positionieren des Schlittens auch bei geschlossenem Deckel noch möglich. Wird der Schlitten nicht ganz nach rechts geschoben, so ist der Abtastvorgang entsprechend schneller.

7.4. Das Einstellen der Helligkeit

Als einziges braucht die Helligkeit eingestellt zu werden. Sie kann viel zur Erkennungsqualität beitragen.

Beim Hawk CP14 wird die Helligkeit mit dem Blauen Regler ("Copy Intensity") am Scanner eingestellt. Richtung "Light" (kleine Zahlen) werden die Bilder heller, richtung "Dark" (grosse Zahlen) werden sie dunkler. Grundsätzlich liefert die Einstellung 3 (mittlere Helligkeit) am Scanner meistens gute Resultate.

Beim Hawk 432 geschieht die Helligkeitseinstellung per Software. Wählen Sie den Menüpunkt "Scanner..." im Menu "Diverses"

an (Abb. 6.1). Unter dem Titel "Helligkeit:" können sie die Helligkeit in 6 Stufen wählen. Kleine Zahlen bedeuten hellere, grosse Zahlen dunklere Bilder.

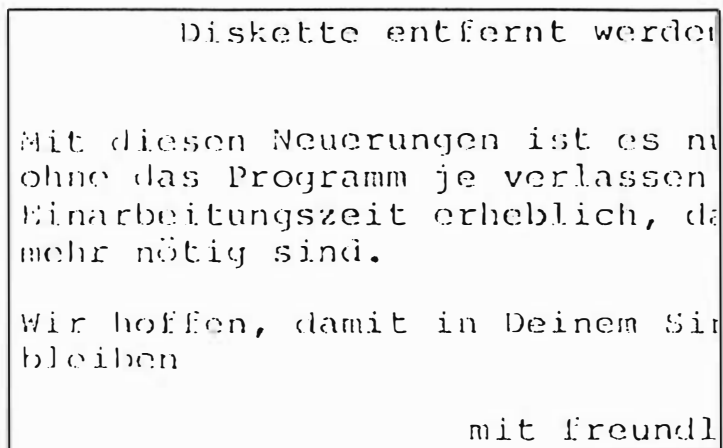


Abb. 7.1. a) Es wurde zu hell gescannt. Die Buchstaben beginnen auseinanderzufallen und werden von AUGUR in mehrere Teile getrennt.

Der Name Helligkeit ist nicht ganz korrekt. Es wird nicht die Gesamthelligkeit des abgetasteten Bildes eingestellt (für die Schrifterkennung werden ja keine Grauwerte gescannt), sondern

die Entscheidungsschwelle zwischen Schwarz und Weiss wird verschoben. Es leuchtet ein, dass bei kontrastarmen Schriften (Kanten gehen von satt Schwarz erst über Grau ins Weiss über) mit dem Helligkeitsregler die Dicke der gescannten Buchstaben variiert werden kann. Dies ist ein entscheidender Faktor für die Schrifterkennung.

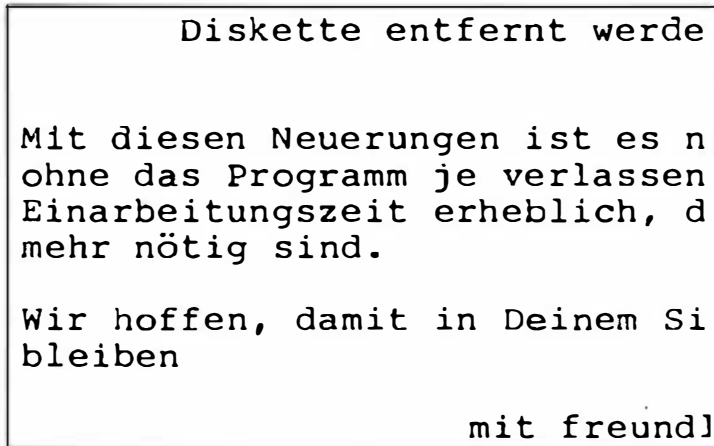


Abb. 7.1. b) Korrekte Einstellung der Helligkeit

Bei zu dünnen Strichen besteht die Gefahr, dass die Buchstaben oben und unten nicht ganz geschlossen werden und Sie AUGUR demzu-

folge für zwei einzelne Buchstaben hält. Bei zu dunkler Einstellung kann es vorkommen, dass die Buchstaben zusammenwachsen und AUGUR sie nicht mehr zu trennen vermag. Diese Problematik wird in den Bildern 7.1 a) - c) illustriert.

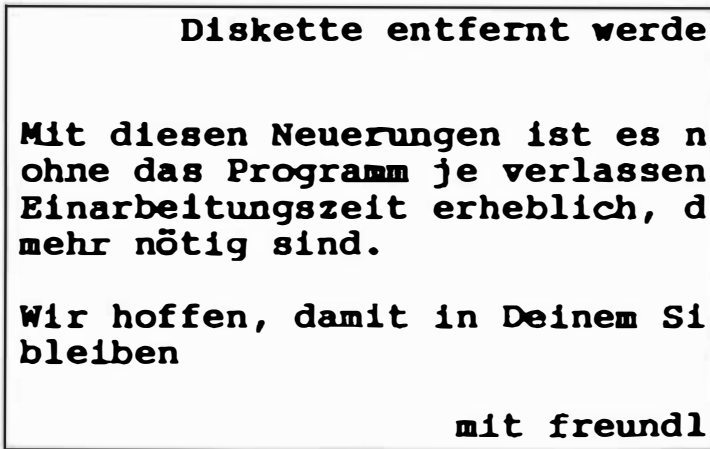


Abb. 7.1. c) Es wurde zu dunkel gescannt. Die Buchstaben beginnen zu verwachsen und können von AUGUR nicht mehr getrennt werden.

**Des weiteren hat die Helligkeit auch einen
Einfluss bei der Verwendung von Bibliotheken**

(siehe Kapitel 11. "Bibliotheken"). Wird eine Schriftart bei einer gewissen Helligkeit gelernt und wird das selbe Dokument bei anderen Helligkeiten noch einmal erkannt, so können, bedingt durch die unterschiedlichen Strichdicken, etliche unbekannte Buchstaben auftreten.

Nehmen Sie noch einmal das Dokument aus dem letzten Kapitel ("Brief") und Scannen Sie dieses bei verschiedenen Helligkeiten. Laden Sie jedes Mal wieder neu die Bibliothek "BRIEF200.LIB" und starten Sie den Erkennungsvorgang. Es wird bei jeder Helligkeitsstufe eine unterschiedliche Anzahl von unbekanntenen Zeichen auftauchen.

7.5. Das Einstellen der Auflösung

Beim Hawk CP14 kann die Auflösung nicht eingestellt werden. Sie beträgt fix 200 Dpi (8 Punkte pro Millimeter).

Beim Hawk 432 kann Sie in den Stufen 200, 300 und 400 Dpi (entspricht 8, 12 und 16 Punkte pro Millimeter) verändert werden. Dies geschieht softwaremässig mit dem Menueintrag "Scanner..." unter dem Menu "Diverses". Es

erscheint das bereits bekannte Dialogfeld (Abb. 6.1), mit der auch die Helligkeit des Hawk 432 eingestellt wird. Unter dem Titel "Auflösung:" finden sich die drei Felder "200", "300" und "400". Die aktuelle Auflösung ist jeweils invertiert (weiss auf schwarz) dargestellt. Durch Klicken in das gewünschte Feld kann die Auflösung verändert werden.

Bei Scannern mit IDC-Treibern geschieht diese Einstellung meistens über das entsprechende IDC-Accessory. Lesen Sie dazu die dem jeweiligen Treiber beigelegten Blättern durch.

7.6. Zur Beachtung

AUGUR besitzt andere Augen als der Mensch.

Damit ist gemeint, dass AUGUR gewisse Dokumente, welche Sie für problematisch halten, ohne Schwierigkeiten lesen kann, und andere, an sich tadellose Dokumente ihm Mühe bereiten.

Wichtig für AUGUR ist, dass gleiche Buchstaben immer gleiches Aussehen haben und dass die einzelnen Buchstaben durch weisse Linien voneinander trennbar, also nicht verwachsen

sind. Das Aussehen der Buchstaben selbst spielt jedoch absolut keine Rolle. So wurde mit AUGUR schon griechische und sogar japanische Schrift erkannt.

7.7. Der IDC-Standard

IDC bedeutet **I**ntelligent **D**river **C**oncept. Die Funktionsweise ist einfach: ein Treiberprogramm übernimmt die Ansteuerung des "fremden" Scanners und AUGUR kommuniziert mit diesem Treiber über eine definierte Software-Schnittstelle. Dieses Treiberprogramm ist beim IDC ein Desk-Accessory. Über dieses Accessory wird der Scanner eingestellt. Der Benutzer kann weiterhin vom Programm AUGUR aus Scannen. AUGUR sendet eine Meldung an das IDC-Accessory und dieses löst den eigentlichen Scanvorgang aus. Diese intensive Kommunikation zwischen AUGUR und IDC-Accessory bleibt für den Benutzer unsichtbar.

Zu allen Scannern der Marvin AG existieren geeignete IDC-Treiber. Natürlich muss es sich bei IDC nicht zwingend um einen physikalisch existierenden Scanner handeln. In Kürze soll ein Treiber erhältlich sein, der Bilder im IMG-

Format von Diskette oder Harddisk lesen kann. Damit können auch solche Bilder von AUGUR verarbeitet werden.

Hier noch ein Hinweis für jene Anwender, welche gleichzeitig mehrere Scanner verwenden möchten. AUGUR sucht beim Aufstarten nach einem vorhandenen Scanner und zwar in der Reihenfolge IDC-Accessory, Hawk 432, Hawk CP14. Der erste gefundene Scanner gewählt. Beim Auslösen eines Scanvorganges von AUGUR aus wird dann immer dieser Scanner angesteuert. Mit dem Dialogfeld "Scannen..." im Menu "Diverses" (Abb. 6.1) kann zwischen den Scannern Hawk CP14 und CP432 umgeschaltet werden. Ist gerade ein IDC-Treiber gewählt, so ist keiner der beiden Scanner selektiert. Beim Klicken auf eines der beiden Felder ("Hawk CP14" oder "Hawk 432") wird auf den entsprechenden Scanner umgeschaltet. Soll auf einen der IDC-Treiber zurückgeschaltet werden, so muss ein Scanvorgang über das entsprechende Accessory ausgelöst werden. AUGUR schaltete dann auf diesen Treiber um. Im Folgenden kann also wieder von AUGUR aus gescannt werden.

8. Bilder speichern und laden

Gescannte Bilder können in Grafikform auf Diskette oder Harddisk gespeichert und von dort wieder geladen werden. So ist es zum Beispiel möglich, ein mehrseitiges Dokument zuerst in den Computer einzulesen, abzuspeichern und dann Bild für Bild zu erkennen.

Bilder werden mit dem Menüpunkt "Bild sichern..." im Menu "Datei" gespeichert. Wird dieser Punkt angewählt, so erscheint die Objektauswahl-Box, in der nach einem Namen für das Bild gefragt wird. Sie können über die Tastatur einen neuen Namen eingeben oder mit der Maus einen bereits bestehenden anwählen. Bilder tragen zur Kennzeichnung den Extender (Erweiterung) ".SCN". Dieser kann über die Tastatur eingegeben oder auch weggelassen werden. AUGUR fügt ihn dann selbständig zum Namen. Mit dem Knopf "ABBRUCH" kann die Objektauswahl-Box wieder verlassen werden, ohne dass etwas verändert wird. Mit "OK" wird das Bild schliesslich gespeichert.

Ist der gewählte Name bereits vorhanden, so erscheint eine Warnbox. Sie können dann

wählen, ob das vorhandene Bild durch das neue ersetzt werden, oder der Speichervorgang abgebrochen werden soll. Im zweiten Fall wird das neue Bild nicht gespeichert.

Gespeicherte Bilder wieder laden können Sie mit dem Menüpunkt "Bild laden..." ebenfalls im Menu "Datei". Es erscheint wieder die Objektauswahl-Box. Wählen Sie mit der Maus das Bild aus, welches Sie laden möchten, und klicken Sie anschliessend auf das Feld "OK". Das Bild wird jetzt in den Computer geladen und anschliessend angezeigt.

Bilder werden immer in einem komprimierten Format gespeichert. Ein unkomprimiertes Bild hat eine Datenmenge von 500 KByte bis 2 MByte, je nach Auflösung des Scanners. Von solchen Bildern hätte also im besten Fall eines pro Diskette Platz. Hingegen kann man 5 bis 20 komprimierte Bilder auf einer Diskette speichern.

Aus Kompatibilitätsgründen zu anderen Programmen, ist AUGUR auch fähig, unkomprimierte Bilder zu laden. Solche Bilder müssen headerlos sein und die fixe Breite von 1664 Pixeln (208 Byte) haben. Die Länge kann variabel sein, jedoch nicht grösser als 2208 Zeilen.

9. Rahmen

Ist ein Bild einmal geladen oder gescannt, so ist der nächste Schritt AUGUR mitzuteilen, welche Teile vom Bild erkannt werden sollen. Dazu können Sie die gewünschten Textstellen mit der Maus einrahmen.

9.1. Rahmen einzeichnen

Rahmen werden immer im verkleinerten Bild eingezeichnet und zwar wie folgt: Bewegen Sie den Mauszeiger an die linke obere Ecke des gewünschten Bereichs. Halten Sie die linke Maustaste gedrückt, während Sie mit der Maus nach rechts unten fahren. Dabei zeichnen Sie ein Rechteck, welches sich wie ein Gummiband (Rubberbox) vergrößert. Der Mauszeiger wechselt dabei in einen zeigenden Finger. Sobald Sie die Maustaste wieder loslassen, bleibt das Rechteck stehen und wird mit einer festen Umrandung gezeichnet (Abb. 9.1).

Wechseln Sie jetzt in den Scroll-Modus, indem Sie den Mauszeiger in die Nähe einer Rahmen-ecke bewegen und die linke Maustaste kurz drücken. Sie sehen das Bild und auch den ein-

gezeichneten Rahmen vergrößert (Abb. 9.2).



Abb. 9.1. Aufziehen eines Rahmens und fertig eingezeichneter Rahmen

Mit den **Cursorpfeilen** (auf der Tastatur) haben Sie die Möglichkeit die exakte Positionierung der linken oberen Rahmenecke noch zu verändern. Drücken Sie [Shift] und gleichzeitig einen der Cursorpfeile, so verändert sich die rechte untere Rahmenecke. Und die Kombination [Alternate] und ein Cursorpfeil bewirkt, dass sich der Bildausschnitt in die jeweilige Richtung bewegt (gleich wie beim Bewegen der Maus).



Abb. 9.2. Rahmenecke im Scroll-Modus

Wenn Sie den Scroll-Modus wieder verlassen (Tastendruck der Maus oder der Tastatur), können Sie mit der Maus einen weiteren Rahmen einzeichnen. Nur der jeweils zuletzt gezeichnete Rahmen kann im Scroll-Modus feinjustiert werden. Links auf dem Bildschirm, unter dem Titel "Info" wird angezeigt, wieviele Rahmen bereits eingezeichnet worden sind. Mit dem Menüeintrag "**Rahmen löschen**" im Menü "Rahmen" oder mit der Taste [Esc] können die Rahmen in umgekehrter Reihenfolge, wie sie gezeichnet wurden wieder gelöscht werden. Mit dem Menüeintrag

"alle Rahmen löschen" werden alle eingezeichneten Rahmen gelöscht.

Mit dem Menüpunkt **"Wissenswertes..."** im Menü "Rahmen" kann eine Zusammenfassung über den Gebrauch von Rahmen angezeigt werden.

9.2. Schrifterkennung und Rahmen

Achten Sie darauf, dass die Rahmengrenzen oben und unten nicht zu eng um den Text gelegt sind. AUGUR braucht dort etwas Platz um die Schräglage des Textausschnittes zu bestimmen. Ist dieser Platz sehr knapp, kann es vorkommen, dass die erste oder die letzte Textzeile nicht erkannt wird. Es stört hingegen nicht, wenn Sie die Zeile darüber oder darunter mit dem Rahmen anschneiden. AUGUR lässt angeschnittene Zeilen immer weg.

Die einzelnen Rahmen werden in der Reihenfolge bearbeitet, in welcher sie eingezeichnet worden sind.

Normalerweise werden die eingezeichneten Rahmen beim Laden oder Scannen eines neuen Dokumentes automatisch gelöscht. Dies kann

unter Umständen hinderlich sein, wenn Sie von verschiedenen Dokumenten mit gleichem Aufbau, immer die selben Ausschnitte erkennen möchten. In diesem Fall wählen Sie den Menüpunkt "**Rahmen erhalten**" im Menü "Rahmen" an. Dieser Menüpunkt wird danach mit einem Haken versehen, was bedeutet, dass die Rahmen von jetzt an beim Laden und Scannen von Bildern nicht mehr automatisch gelöscht werden.

Wenn Sie keine Rahmen eingezeichnet, oder die bestehenden wieder gelöscht haben, erkennt AUGUR die gesamte Vorlage. In Wirklichkeit sind nie alle Rahmen gelöscht, es verbleibt mindestens ein Rahmen, der das gesamte Dokument umfasst.

9.3. Rahmenmarker

Normalerweise wird zwischen den Text aus zwei Rahmen eine Leerzeile gefügt. Dies reicht aus, um die Textblöcke optisch zu trennen. Manchmal soll der erkannte Text aber von Programmen weiterverarbeitet werden. Ein solches Programm könnte zum Beispiel die Speicherung des Textes in eine Datenbank übernehmen. Dazu ist es notwendig den Text

aus den verschiedenen Rahmen sicher trennen zu können. Eine Leerzeile allein erfüllt diese Aufgabe nicht.

Aus diesem Grund wurden die Rahmenmarker geschaffen. Ist diese Option (Eintrag "**Rahmenmarker**" im Menu "Text") angeklickt (der Eintrag wird darauf mit einem Haken versehen), so wird beim Speichern des erkannten Textes oder beim Senden über die RS232 Schnittstelle, zwischen den Text aus jedem Rahmen ein Formfeed-Zeichen gesetzt (12 oder in Hex C). Ein entsprechendes Programm kann damit den Text aus verschiedenen Rahmen sicher trennen. Im weiteren wird auch zwischen zwei erkannte Seiten (mit "Text anhängen" in der gleichen Datei gespeichert) ein Formfeed gesetzt. Durch erneutes Anklicken von "Rahmenmarker" wird diese Option wieder ausgeschaltet.

10. Der Erkennungsmodus

Im Menu "Schrifterkennung" finden Sie drei verschiedene Erkennungsmodi: "**schneller Modus**", "**goldenes Mittel**" und "**präziser Modus**".

Hiermit kann gewählt werden, wie genau AUGUR die Vorlage untersuchen soll. Beim präzisen Modus werden wenige Fehler gemacht, die Rechenzeit ist aber auch bedeutend höher als bei den anderen Modi. Es kann also in gewissen Grenzen Geschwindigkeit gegen Genauigkeit ausgetauscht werden. Dies lässt sich leicht nachprüfen, wenn Sie eines der mitgelieferten Beispieldokumente hintereinander mit den drei verschiedenen Modi erkennen.

10.1. Wahl des richtigen Erkennungsmodus

Ausschlaggebend für eine gute Wahl des Erkennungsmodus sind vier Kriterien: Die Schriftgröße, die Strichdicke, die Schärfe der Schrift und die Unterschiedlichkeit der verschiedenen Buchstaben. Die ersten beiden

Kriterien werden natürlich durch die Auflösung des Scanners entscheidend beeinflusst.

Wenn Sie mit 400 Dpi scannen (Hawk 432), ist die Sache einfach. Sie können dann in den meisten Fällen den schnellen Modus verwenden und damit sehr gute Resultate erzielen. Die unten gegebenen Hinweise haben ihre Gültigkeit nur bei kleiner Auflösung (200, 300 Dpi). Trotzdem ist es auf jeden Fall von Vorteil, diese Hinweise zu studieren und mit verschiedenen Vorlagen ein wenig zu experimentieren.

Die **Schriftgrösse** ist erstaunlicherweise nicht das wichtigste Kriterium. Tendenziell gilt aber sicher, dass für grössere Schriften ein schnellerer Modus gewählt werden kann, als für kleinere.

Wichtiger ist die **Strichdicke** der Buchstaben. Für "dünne" Schriften muss der präzise Modus gewählt werden. Man kann sich das so vorstellen, dass AUGUR im schnellen Modus die Vorlage durch ein unscharfes Auge betrachtet. Je präziser der gewählte Modus, desto schärfer stellt sich dieses Auge ein. Dünne Striche, auch wenn sie von grossen Buchstaben stammen, können im schnellen Modus zum Teil gar nicht gesehen werden.

Mit **Schärfe** ist die Übereinstimmung gleicher Buchstaben der selben Schriftart gemeint. Dies ist bei gedruckten Vorlagen meistens optimal der Fall. Schwieriger wird es bei Dokumenten von Matrixdruckern, bei unterschiedlich hart angeschlagener Schreibmaschine oder bei wechselnd schwärzendem Farbband (die Buchstaben werden unterschiedlich dick). Der gleiche Fall tritt auch bei einer Fotokopie mit unregelmässiger Schwärzung auf. Bei solchen Vorlagen sollte möglichst nicht der präzise Modus gewählt werden. AUGUR würde wahrscheinlich gleiche Buchstaben mit unterschiedlicher Schwärzung für verschieden halten und damit die Bibliothek stark anwachsen lassen.

Letztes wichtiges Kriterium ist die **Unterschiedlichkeit** der verschiedenen Buchstaben. Bei gewissen Schriften ist zum Beispiel das grosse "I" und das kleine "l" sehr ähnlich oder sogar genau gleich. Weitere Beispiele wären die Buchstabenpaare "f" und "t", die Null und das "O" oder das "a", das "e" und das "o". Bei solchen Schriften sollte der präzise Modus gewählt werden. Es ist jedoch Ermessenssache, ob man eine längere Erkennungszeit und eine grössere Bibliothek für eine kleinere Fehlerquote in Kauf nehmen soll. In einem Textverarbeitungsprogramm können solche Verwechslungen von

Buchstabenpaaren mit der Funktion "Suchen und Ersetzen" schnell behoben werden (siehe auch Kapitel 16 "Tips zum Gebrauch von Textverarbeitungsprogrammen").

10.2. Faustregeln für die Wahl des Erkennungsmodus

Exakte Werte für die obigen vier Kriterien können leider nicht gegeben werden, da vor allem ihre Kombination den besten Erkennungsmodus bestimmt. Oft ist es auch Ermessenssache des Benutzers. Tendenziell lässt sich sagen, dass bei präziseren Modi weniger Fehler auftauchen, dies jedoch mit höherer Erkennungszeit und grösseren Bibliotheken erkauft werden muss.

Im folgenden seien einige Faustregeln für die Wahl der drei Erkennungsmodi gegeben:

schneller Modus: AUGUR liest in diesem Modus schnell, jedoch nicht so präzise. Er ist bei grossen Schriften oder bei klassischer Schreibmaschinenschrift (Courier) angebracht. Auch bei etwas "verzitterten" Schriften, wie zum Beispiel Matrix-druckerschrift oder einer schlechten

Kopie, führt dieser Modus meistens zu besseren Resultaten, als der präzise Modus (siehe "präziser Modus"). Bei hoher Auflösung des Scanners (400 Dpi) sollte dieser Modus auch für Zeitungen und Bücher verwendet werden.

goldenes Mittel: Dieser Modus kann für moderne und kleine Schreibmaschinenschriften verwendet werden. Ebenso ist er brauchbar für gross geschriebene Drucksachen (ab ca. 3 mm Buchstabenhöhe oder 9 Punkten).

präziser Modus: In diesem Modus arbeitet AUGUR sehr genau und ist dementsprechend etwas langsamer. Er ist bei klein geschriebenen Drucksachen (wie z.B. Zeitung) angebracht. Vorsichtshalber kann immer mit diesem Modus begonnen werden. Bei unscharfen oder verzitterten Schriften, sowie bei ungleichmässig belichteten Kopien kann es jedoch vorkommen, dass AUGUR gleiche Buchstaben für verschieden hält und die Bibliothek stark anwächst. In diesem Fall sollte auf einen schnelleren Modus umgeschaltet werden.

11. Bibliotheken

Eine Bibliothek beinhaltet Prototypen der verschiedenen Zeichen einer bestimmten Schriftart. AUGUR vergleicht beim Erkennungsvorgang jedes Zeichen mit diesen Prototypen. Wird ein passender (genügend ähnlicher) Prototyp gefunden, so gilt das Zeichen als erkannt und wird in Textform gespeichert. Kann kein genug ähnliches Symbol gefunden werden, so gilt das Zeichen als neu und wird der Bibliothek hinzugefügt. Die Bibliothek erweitert sich so automatisch. Jeweils am Schluss des Erkennungsvorganges werden diese neuen Zeichen dem Benutzer präsentiert und er kann über die Tastatur eingeben um welches Zeichen es sich handelt. Anschliessend werden die bis dahin unbekanntenen Zeichen ebenfalls in Textform übersetzt.

Die gewünschte Bibliothek muss zuerst in den Arbeitsspeicher des Computers geladen werden. Es kann sich zur Zeit nur eine Bibliothek im Arbeitsspeicher befinden. Einmal erstellte oder erweiterte Bibliotheken können für den späteren Gebrauch auf Diskette oder Harddisk gespeichert werden.

Der Name der gerade aktiven (also im Arbeits-

speicher befindenden) Bibliothek wird links auf dem Bildschirm unter dem Titel "Bibliothek" angezeigt. Trägt die Bibliothek keinen Namen (weil sie noch nie gespeichert wurde) oder befindet sich keine Bibliothek im Arbeitsspeicher, so bleibt das entsprechende Feld leer. Wurde eine Bibliothek während des Erkennungsvorganges erweitert, so erscheint rechts von diesem Feld das Wort "neu" um anzuzeigen, dass die Bibliothek im Arbeitsspeicher nicht mehr mit jener auf der Diskette übereinstimmt. Speichern Sie die Bibliothek erneut, so verschwindet das Wort "neu".

Ebenfalls unter dem Titel "Bibliothek" wird angezeigt wieviele Prototypen (Zeichen) die aktive Bibliothek enthält.

11.1. Bibliotheken laden

In den Arbeitsspeicher geladen werden Bibliotheken mit dem Menüpunkt "**Bibliothek laden...**" im Menu "Datei". Wird dieser Menueintrag angewählt, so erscheint die Objektauswahl-Box in der die gewünschte Bibliothek ausgewählt werden kann. Es werden dabei nur Bibliotheken, also Dateien mit dem Extender ".LIB", angezeigt. Nach Anwählen

des "OK"-Feldes wird die zuvor selektierte Bibliothek geladen und steht Ihnen fortan im Arbeitsspeicher des Computers zur Verfügung. Des weiteren wird der Name links auf dem Bildschirm unter dem Titel "Bibliothek" eingetragen.

11.2. Bibliotheken sichern

Erweiterte oder neu erstellte Bibliotheken können mit dem Menüpunkt "Bibliothek sichern..." im Menu "Datei" auf Diskette geschrieben werden. Nach Anwählen dieses Menüpunktes erscheint die Objektauswahl-Box, in welcher ein neuer Name eingegeben, oder eine bereits existierende Bibliothek ausgewählt werden kann. Es wird jeweils der Name der zuletzt geladenen oder gespeicherten Bibliothek vorgeschlagen. War keine Bibliothek geladen (wurde die Bibliothek also vollständig neu aufgebaut), schlägt AUGUR den Namen des aktuellen Bildes vor (Anstelle des Extenders ".SCN" steht natürlich der Extender ".LIB"). Trägt dieses keinen Namen, so bleibt die entsprechende Zeile in der Objektauswahl-Box leer.

Wählen Sie einen bereits existierenden Namen,

so wird die alte Datei durch die neue ersetzt. Die alte Bibliothek geht also verloren. AUGUR warnt Sie jedoch vor der Ausführung dieses Schrittes mit einer Warnbox. Sie können dann wählen, ob die alte Datei überschrieben oder ob der Speichervorgang abgebrochen werden soll.

Mit dem Speichern einer Bibliothek wird der gewählte Name auch in das Feld unter dem Titel "Bibliothek" auf dem Bildschirm übernommen. Zusammen mit der Bibliothek werden folgende Daten abgespeichert: Der Erkennungsmodus, die Auflösung (beim Hawk 432) und ob folgende Optionen eingeschaltet sind: "kein Einrücken", "Fixed Space", "Linien immer Entf.". Dabei gelten die aktuellen Einstellungen, nicht also jene, zum Zeitpunkt des Scannens. Beim späteren Laden der Bibliothek werden diese Parameter wieder richtig eingestellt.

11.3. Bibliothek löschen

Damit ist das Löschen der Bibliothek im Arbeitsspeicher gemeint. Auf Diskette gespeicherte Bibliotheken können ebenfalls gelöscht werden. Lesen Sie dazu im Kapitel 15 "Wei-

tere Funktionen" nach.

Die Bibliothek im Arbeitsspeicher wird mit dem Menüpunkt "**Bibliothek löschen**" im Menü "Datei" gelöscht. Wird jetzt ein Text erkannt, so baut sich eine vollständig neue Bibliothek auf. Achten Sie darauf, dass vor dem Löschen keine Sicherheitsabfrage vorgenommen wird. Erweiterte oder neu aufgebaute Bibliotheken gehen also verloren, falls sie noch nicht auf Diskette gespeichert worden sind (erkennbar am Wort "neu" unter dem Titel "Bibliothek" auf dem Bildschirm).

Das Löschen der Bibliothek ist dann sinnvoll, wenn eine unbekannte Schrift erkannt werden soll. Eine bereits geladene Bibliothek könnte dann nichts zum Erkennungsvorgang beitragen und würde diesen nur verlangsamen.

11.4.

Das Wiederverwenden gespeicherter Bibliotheken

Prinzipiell braucht AUGUR jede Schriftart nur einmal beigebracht zu werden. Falls sie abgespeichert wurde, steht sie von da an jederzeit wieder zur Verfügung. Dazu ist jedoch, vor

allem beim Arbeiten mit dem Hawk CP14, einige Vorsicht geboten:

Schwierigkeiten können vor allem bei Verwendung des präzisen Modus auftreten. Dieser Modus duldet nur sehr kleine Abweichungen in Form und Dicke der einzelnen Buchstaben. Konsequenz bei zu stark abweichender Schriftart ist ein starkes Anwachsen der Bibliothek und eine damit verbundene Verlangsamung des Erkennungsprozesses. Weitaus weniger kritisch ist die Sache bei Verwendung des mittleren oder des schnellen Modus. Das im folgenden Gesagte gilt deshalb vor allem für den präzisen Modus. Beim Arbeiten mit hoher Auflösung (400 Dpi) treten deshalb normalerweise keine Schwierigkeiten auf, weil dabei ja meist mit dem schnellen Modus gearbeitet wird.

Dokumente mit scharfer Schrift, wie von Druckern mit Carbonfarbband, Hochglanzprospekte oder Zeitschriften, bieten normalerweise keine Schwierigkeiten. Vorsicht ist geboten bei Zeitungen mit schlechtem Kontrast oder mehrfachen Kopien, also bei unscharfen Schriften. Bei solchen Schriften hat eine kleine Verstellung der Helligkeit am Scanner meistens eine Auswirkung auf die Buchstabendicke.

Wird ein mehrseitiges Dokument mit unscharfer Schrift bei gleichbleibender Scannereinstellung erkannt, so treten keine Schwierigkeiten auf. Wird jedoch Tage darauf mit einer anderen Einstellung weitergefahren, kann es sein, dass die bis dahin erstellte Bibliothek fast nicht mehr zu gebrauchen ist.

Im Zweifelsfalle lohnt es sich die alte Bibliothek einmal zu verwenden. Ist der Erkennungsprozess nach einigen Zeilen noch deutlich langsamer als gewohnt und sieht man, dass sich die Zahl der Prototypen stark vergrössert, so ist es besser, den Erkennungsvorgang abbrechen (Tastendruck) und die Bibliothek aus dem Arbeitsspeicher zu entfernen.

Meistens bietet AUGUR auch dann noch grossen Nutzen, wenn die Bibliothek neu aufgebaut werden muss. Eine DIN A4 Seite beispielsweise beinhaltet leicht 3000 Schriftzeichen. Die zugehörige Bibliothek ist vielleicht 50 bis 100 Zeichen gross. Das heisst, bei Verwendung von AUGUR haben Sie 30 bis 60 Mal weniger Tipp-Arbeit.

Möchte man für eine gewisse Schrift eine Bibliothek aufbauen, welche für möglichst viele verschiedene Dokumente anwendbar ist, so lohnt sich meistens folgendes Vorgehen: Neh-

men Sie eine saubere Vorlage mit der entsprechenden Schrift und erkennen Sie diese mehrmals, wobei sie die Helligkeit am Scanner absichtlich jedesmal etwas verstellen. Damit erhalten sie Prototypen der Buchstaben in verschiedenen Dicken. Die so erstellte Bibliothek speichern sie weitere Male möglichst nur noch dann ab, wenn saubere Buchstaben hinzugekommen sind. Eine Bibliothek ständig zu erweitern kann die Folge haben, dass zum Beispiel ein dickes "i" für ein "t" gehalten wird oder ein angekratztes "e" für ein "c". Mit anderen Worten: Die Wahrscheinlichkeit für Fehler steigt.

11.5. Lange Dokumente

Auch nach vielen Seiten der gleichen Schriftart können für AUGUR immer noch neue Zeichen auftreten. Es handelt sich dabei meistens um verunreinigte oder nur halb gedruckte Zeichen. Seien Sie vorsichtig mit dem ständigen Erweitern der Bibliothek.

Ein guter Weg um mit langen Dokumenten zu arbeiten ist eine Bibliothek über wenige Seiten aufzubauen und abzuspeichern. Kommen von diesem Zeitpunkt an neue Zeichen hinzu, so

speichern Sie die erweiterte Bibliothek nicht mehr ab. Wenn der Erkennungsprozess wegen einer grossen Bibliothek im Arbeitsspeicher deutlich langsamer geworden ist, laden Sie die "alte" Bibliothek von der Diskette wieder. Auf diese Weise wird ein flüssiges Arbeiten mit AUGUR garantiert und ein Überlaufen der Bibliothek verhindert.

12. Das Starten der Schrifterkennung

Voraussetzung für das Starten der Schrifterkennung ist, dass sich ein Bild im Arbeitsspeicher befindet und dass die Bibliothek gelöscht oder die richtige Bibliotheks-Datei geladen ist. Zusätzlich ist darauf zu achten, dass der richtige Erkennungsmodus eingeschaltet ist (siehe Kapitel 10 "Der Erkennungsmodus") und falls gewünscht die Rahmen richtig eingezeichnet sind (Kapitel 9 "Rahmen").

Gestartet wird die Schrifterkennung mit dem Menüpunkt "Text erkennen" im Menu "Schrifterkennung". Befindet sich bereits ein Text im Arbeitsspeicher des Atari, so erscheint eine Warnbox, welche Ihnen die Wahl zwischen dem Löschen des alten Textes, dem Anhängen des neuen Textes zum Alten und dem Abbrechen des Erkennungsvorganges lässt.

Nach dem Start wird Rahmen für Rahmen abgearbeitet. Der Erkennungsvorgang für einen einzelnen Rahmen unterteilt sich in zwei Schritte: Das Segmentieren und das eigentliche Erkennen. Anschliessend wird automatisch in den Lernmodus (nur falls im Text Zeichen waren, die AUGUR noch nicht kennt) und an-

schliessend in den Textanzeige-Modus verzweigen. Der Lernmodus wird im nächsten und der Textanzeige-Modus wird im übernächsten Kapitel beschrieben.

Unter dem Titel "Arbeit" (links auf dem Bildschirm) wird angezeigt, welcher Rahmen von AUGUR gerade bearbeitet wird, unter dem Titel "Info" steht die Anzahl insgesamt eingezeichneter Rahmen.

12.1. Das Segmentieren

Das Segmentieren einer ganzen A4 Seite dauert 10 bis 30 Sekunden. Bei kleineren Blöcken ist AUGUR entsprechend schneller. Dieser Vorgang wird angezeigt durch das Wort "Segmentieren" unter dem Titel "Arbeit". Wegen der kleineren Datenmenge sind Bilder, welche mit niedriger Auflösung gescannt wurden, schneller segmentiert, als solche, welche mit hoher Auflösung gescannt worden sind.

Während dem Segmentieren untersucht Augur das Dokument auf seine Schräglage hin und sucht anschliessend den Ort und die Grösse der einzelnen Schriftzeilen. Eine nicht allzu grosse Abweichung von der Horizontalen (ca. 5 mm

bezogen auf die Breite einer A4 Seite) wird toleriert. Auch müssen nicht alle Schriftzeilen zueinander parallel liegen (ein Dokument, welches mehrmals in die Schreibmaschine eingespannt wurde, oder eine mit Schere und Leim editierte Vorlage), AUGUR untersucht die Schräglage für jede Schriftzeile einzeln.

Wird während der Segmentierung irgend eine Taste gedrückt, so wird diese noch zu Ende geführt und der Erkennungsvorgang danach abgebrochen.

12.2. Der eigentliche Erkennungsvorgang

Nach der Segmentierung wird das Wort "Segmentieren" vom Bildschirm entfernt und an jener Stelle angezeigt, wieviele Schriftzeilen sich im entsprechenden Block befinden und an welcher Zeile AUGUR zur Zeit arbeitet. Unter dem Titel "Info" steht zusätzlich, wieviele Zeichen bereits gelesen wurden. Diese Zahl wird aus Zeitgründen nur am Ende jeder Schriftzeile nachgeführt.

Der Erkennungsvorgang kann durch Drücken irgend einer Taste unterbrochen werden.

AUGUR beendet noch die aktuelle Zeile und bricht sodann ab. Es besteht danach allerdings keine Möglichkeit mehr, den Erkennungsvorgang an der selben Stelle wieder aufzunehmen.

13. Der Lernmodus

Im Lernmodus kann der Benutzer interaktiv jedem Schriftzeichen aus der Bibliothek den richtigen Buchstaben zuweisen. AUGUR präsentiert dazu das bis dahin noch unbekannte Zeichen in grafischer Form. Nachdem der zugehörige Buchstabe über die Tastatur eingegeben wurde, erscheint das nächste Zeichen. Bei diesem Vorgang übersetzt AUGUR das zuvor erkannte Dokument und lernt gleichzeitig die eingegebenen Zeichen.

13.1. Aufruf des Lernmodus

Der Lernmodus wird durch Anklicken des Menüpunktes "Schrift lernen" im Menu "Schrifterkennung" aufgerufen. Mit der [Esc]-Taste kann er wieder verlassen werden. Weiter wird, nachdem die Schrifterkennung gestartet und das Dokument fertig erkannt wurde, automatisch in den Lernmodus verzweigen. Abb. 13.1 zeigt den Bildschirm, wie er sich im Lernmodus präsentiert.

Auf dem Bildschirm befinden sich insgesamt fünf Fenster. Im mittleren quadratischen Fen-

ster ist vergrößert die Grafik des aktuellen Buchstabens sichtbar.

Im oberen Fenster ist der selbe Buchstabe noch einmal dargestellt (dünn umrahmt). Diesmal allerdings innerhalb des Umfeldes, welchem er entnommen wurde. Dieses Umfeld kann bei der Bestimmung des richtigen Buchstabens zum Teil sehr nützlich sein.



Abb. 13.1. Bildschirmaufbau im Lernmodus

Das Umfeld der Buchstaben kann natürlich nur solange dargestellt werden, wie sich das Originalbild im Speicher befindet. Bei jenen Zei-

chen, welche nicht vom aktuellen Bild stammen, bleibt dieses Fenster deshalb leer.

- Im unteren Fenster wird angezeigt, wieviele Buchstaben die aktuelle Bibliothek insgesamt hat und welche Nummer das aktuelle Zeichen trägt. Auf der unteren Linie steht schliesslich, welcher Buchstabe (es können auch mehrere sein) vom Benutzer dem Zeichen zugeordnet wurde. Ist das Zeichen noch nicht bestimmt, so steht an dieser Stelle ein umgekehrtes Fragezeichen ("¿").

13.2. Das Übersetzen einer Bibliothek

Beim Eintreten in den Lernmodus wird automatisch zum ersten Zeichen gesprungen, dem noch kein Buchstabe zugewiesen ist, das also ein umgekehrtes Fragezeichen hat. Betreten Sie den Lernmodus bei einer Bibliothek, in der bereits alle Zeichen übersetzt wurden, so ertönt eine Glocke und AUGUR springt zum letzten Zeichen. Bei einer solchen Bibliothek braucht natürlich nichts mehr eingetippt zu werden.

Im Anhang B befindet sich eine Zusammenfas-


sung sämtlicher Manipulationsmöglichkeiten im Lernmodus. Diese werden im folgenden detailliert beschrieben.

Dem aktuellen (angezeigten) Zeichen kann der zugehörige Buchstabe durch einfachen Tastendruck zugewiesen werden. AUGUR lernt diesen Buchstaben sofort und springt automatisch zum nächsten Zeichen. Haben Sie sich bei einem Zeichen vertippt, so gelangen sie mit [**Backspace**] oder dem **linken Cursorpfeil** zum vorhergehenden Zeichen zurück. Nach dem Übersetzen des letzten Zeichens ertönt eine Glocke und AUGUR springt nicht mehr weiter.

Tritt ein verwachsenes Zeichen auf (mehrere Buchstaben, die AUGUR nicht zu trennen vermochte), so können Sie diesem mehrere Buchstaben (String) zuweisen. Drücken Sie dazu erst die Taste [**Return**] oder [**Enter**]. Jetzt können Sie der Reihe nach die zugehörigen Buchstaben eingeben (maximal 4). Schliessen Sie die Eingabe wieder mit [**Return**] oder [**Enter**] ab. AUGUR springt dann zum nächsten Zeichen. Falls Sie sich beim Eingeben einer Zeichenkette vertippt haben, können Sie mit [**Backspace**] die eingegebenen Buchstaben der Reihe nach wieder löschen.

Handelt es sich bei einem Zeichen um einen Flecken, eine Büroklammer, oder sonst etwas, dem kein Buchstabe zugewiesen werden soll, so drücken Sie die Leertaste oder **[Delete]**. Der Unterschied besteht darin, dass bei der Leertaste an den entsprechenden Stellen im Text wirklich ein Leerschlag eingefügt wird, bei **[Delete]** hingegen nicht.

Bemerken Sie erst beim Durchlesen des erkannten Textes, dass sich in der Bibliothek vermutlich noch falsche Übersetzungen befinden (ein "b" wurde beispielsweise konstant als "v" übersetzt), so kehren Sie noch einmal in den Lernmodus zurück. AUGUR springt dann ans Ende der Bibliothek. Mit **[Clr Home]** gelangen Sie bei Bedarf an deren Anfang. Der rechte oder der linke **Cursorpfeil** führt zum nächsten oder zum vorhergehenden Buchstaben. Betätigen Sie bei gedrückter **[Shift]**-Taste eine dieser **Cursor-Tasten**, so springt Augur jeweils 10 Zeichen vor oder zurück.

Betätigen Sie während der Eingabe einer Zeichenkette eine dieser Cursor-Tasten, so springt AUGUR ebenfalls zum nächsten oder vorgängigen Zeichen, die gemachten Änderungen an

der Zeichenkette werden dabei allerdings rückgängig gemacht.

Wie bereits erwähnt, kann mit der [Esc]-Taste der Lernmodus wieder verlassen werden.

Wenn Sie bei einer Schrift, die AUGUR bereits kennt, darauf verzichten wollen die neu dazugekommenen Buchstaben zu übersetzen, so können Sie den Lernmodus gleich beim Eintreten mit [Esc] wieder verlassen. AUGUR setzt für jedes nicht übersetzte Zeichen ein umgekehrtes Fragezeichen an die entsprechenden Stellen im Text.

13.3. Sonderzeichen

Abhängig von Tastatur und Sprache, können nicht immer alle gewünschten Buchstaben über die Tastatur erreicht werden. Der deutschen Tastatur fehlen zum Beispiel jegliche Buchstaben mit Akzent.

Solche Sonderzeichen befinden sich im linken Fenster auf dem Bildschirm. Sie können eingegeben werden, indem Sie mit der Maus auf das entsprechende Zeichen fahren und die linke Maustaste betätigen. Als Alternative dazu wurden die Funktionstasten F1 bis F6 mit sogenannten Totzeichen belegt. Die Legende dazu

befindet sich in der untersten Zeile des Fensters mit den Sonderzeichen. Wollen Sie beispielsweise ein "é" eintippen, so drücken Sie zuerst [F1] und anschliessend das e. Dies funktioniert natürlich nur mit jenen Zeichen, welche der Atari auch darstellen kann, also mit all jenen, welche links auf dem Bildschirm angezeigt werden.

Hinweis: Andere Textverarbeitungssysteme als der Atari verwenden zum Teil andere Codes für die Sonderzeichen. An den entsprechenden Textstellen erscheinen dann andere oder gar keine Zeichen. Wenn Sie also einen anderen Computer als den Atari für die Textverarbeitung benutzen, so probieren Sie dies zuerst aus. Oft können Sie auf dem neuen System durch die Funktion "Suchen und Ersetzen" die "falschen" Zeichen gegen die gewünschten austauschen. Falls dies nicht funktioniert müssen Sie auf die Verwendung von Sonderzeichen verzichten. Schreiben Sie im Lernmodus beispielsweise für das "à" ein "a" und korrigieren Sie dies erst im Textverarbeitungsprogramm.

14. Der erkannte Text

AUGUR ist fähig, den erkannten Text auf dem Bildschirm anzuzeigen, ihn unter verschiedenen Formaten auf Diskette abzuspeichern, an bereits bestehende Dateien anzuhängen, oder über die serielle Schnittstelle direkt an ein anderes Computer- oder Textverarbeitungs-System zu senden.

14.1. Den Text anzeigen

Der erste Wunsch besteht natürlich darin, den erkannten Text auf dem Bildschirm sichtbar zu machen. Dies geschieht entweder automatisch nach Starten der Schrifterkennung oder manuell durch anwählen des Menüpunktes "anzeigen" im Menu "Text" (Abb. 14.1).

Um zur nächsten Bildschirmseite "vorzublättern" drücken Sie eine Taste der Tastatur (ausser [Esc]) oder die linke Maustaste. Nach dem Anzeigen der letzten Seite kehrt AUGUR automatisch in das Menu zurück. Sie erkennen die letzte Seite meistens daran, dass der Bildschirm nicht mehr vollständig mit Text gefüllt ist. Der Anzeigemodus kann auch vor der letz-

ten Seite durch Drücken von [Esc] verlassen werden.

Horizontal werden maximal 80 Zeichen dargestellt. Es kann gelegentlich vorkommen, dass erkannte Texte breiter sind. In diesem Fall werden nur die ersten 79 und das letzte Zeichen der Zeilen dargestellt. Es handelt sich dabei jedoch lediglich um ein Anzeigeproblem und bedeutet keinesfalls, dass die restlichen Zeichen verloren sind. Der vollständige Text kann also ganz normal gespeichert oder an einen anderen Computer gesendet werden.

Deinem Wunsche entsprechend haben wir in letzter Minute noch folgende Erweiterungen im Programm AUGUR vorgenommen:

- Die Funktion "Disk formatieren" erlaubt es, eine leere Diskette direkt von AUGUR zu formatieren, ohne das Programm verlassen zu müssen.
- Mit der Funktion "Datei Löschen" können einzelne Dateien, wie Bibliotheken, Texte oder Bilder, wieder von einer Diskette entfernt werden.

Mit diesen Neuerungen ist es nun möglich, mit AUGUR zu arbeiten, ohne das Programm je verlassen zu müssen. Dies verkürzt die Einarbeitungszeit erheblich, da keine Kenntnisse des GEM Desktop mehr nötig sind.

Wir hoffen, damit in Deinem Sinne gehandelt zu haben und verbleiben

mit freundlichen Grüßen:

Urs Müller

Michael Schenkel

Abb. 14.1. Der erkannte Text wird auf dem Bildschirm angezeigt.

Hinweis: Besitzer eines Hawk CP14 können sehr schnell einen Ausdruck des erkannten Textes erhalten. Sie müssen dazu den Text auf Diskette oder Harddisk speichern. Anschliessend kann er mit dem "CPLIST"-Accessory (gehört zum Lieferumfang des Hawk CP14) auf dem Thermodrucker des Hawk CP14 ausgedruckt werden, ohne dass das Programm AUGUR verlassen werden muss.

14.2. Den Text sichern

Erkannte Texte können mit dem Menüpunkt "sichern..." im Menü "Text" auf Diskette oder Harddisk gespeichert werden. Wählen Sie diesen Menüpunkt mit der Maus an, so erscheint die Dateiauswahl-Box, welche Ihnen Gelegenheit gibt einen Namen für den Text einzugeben oder eine bereits vorhandene Datei auszuwählen. Im zweiten Fall wird diese Datei mit dem neuen Text überschrieben, der alte Text geht also verloren! AUGUR unternimmt diesen Schritt jedoch nicht, ohne Sie vorher zu warnen.

14.3.

Den Text einer bestehenden Datei anhängen

Besteht ein Dokument aus mehreren Seiten, so können Sie mit dieser Option die zweite und alle folgenden Seiten an die bereits bestehende Textdatei anhängen. Es resultiert ein zusammenhängender Text. Klicken Sie dazu im Menu "Text" den Eintrag "anhängen..." an. In der darauf erscheinenden Objektauswahl-Box können sie die gewünschte Datei auswählen. Der Name der zuletzt gespeicherten Textdatei wird bereits vorgeschlagen. Ist er korrekt, so können Sie direkt mit der Maus das "OK"-Feld anwählen oder [Return] drücken. Der Text wird dann der entsprechenden Datei angehängt.

Zwischen den bestehenden und den neuen Text werden jeweils drei Leerzeilen eingefügt.

14.4.

Den Text über die serielle Schnittstelle senden

Arbeiten Sie für die Textverarbeitung lieber mit einem anderen Computer, so bietet Ihnen

AUGUR die Möglichkeit, die erkannten Texte über die serielle Schnittstelle (RS232) direkt an einen zweiten Computer zu senden. Sie brauchen dazu ein Modemprogramm für den Empfängercomputer und ein sogenanntes Nullmodem, ein Kabel mit welchem die beiden Computer verbunden werden. Wenden Sie sich dazu an Ihren Händler.

Klicken Sie im Menu "Text" den Eintrag "**senden**" an. Die Datenübermittlung beginnt dann sofort. Sie kann jederzeit mit einem Tastendruck unterbrochen werden.

Die empfangenen Daten können Sie anschließend auf Ihrem gewohnten Programm weiterverarbeiten. Diese Methode bietet daneben den Vorteil, dass AUGUR in dieser Zeit bereits die nächsten Seiten erkennen kann.

Damit die Datenübertragung reibungslos abläuft müssen die Schnittstellen von beiden Computern gleich konfiguriert sein. Für die richtigen Einstellungen im Empfängercomputer lesen Sie bitte im entsprechenden Handbuch nach. Es ist an dieser Stelle unmöglich, die Konfiguration für alle erhältlichen Computer zu beschreiben.

Auf der Seite von Atari steht Ihnen dazu das

Programm "EMULATOR.ACC" zur Verfügung (gehört zum Lieferumfang des Atari Computers und befindet sich ebenfalls auf der mit AUGUR gelieferten Beispiel-Diskette). Achten Sie darauf, dass sich dieses Programm auf Ihrer Bootdiskette befindet (jene Diskette, welche sich beim Einschalten des Computers im Laufwerk A befindet). Das Programm wird beim Aufstarten des Systems geladen und bleibt bis zum Ausschalten resident im Arbeitsspeicher erhalten. In AUGUR erscheinen dann unter dem Menu "Desk" die beiden Einträge "RS 232 Einstellung" und "VT52 Emulator". Bei einer neuen Version dieses Programms erscheint nur noch "VT52 Emulator".

Um die RS232 Schnittstelle Ihres Atari zu konfigurieren klicken Sie falls vorhanden auf den Eintrag "RS 232 Einstellung" oder den Eintrag "VT52 Emulator" und drücken dann die Taste [Help]. In beiden Fällen erscheint ein Dialogfeld, wie in Abb. 14.2 dargestellt.

Stellen Sie die Konfiguration auf dem Atari und dem Empfänger-Computer gleich ein. Die Einstellung auf dem Atari beenden Sie mit dem Feld "OK" oder "ABBRUCH", falls sie die gemachten Änderungen rückgängig machen möchten. Im ersten Fall gelangen Sie sofort zu

AUGUR zurück, im zweiten Fall erst nach Drücken der Taste [Undo]. Bei der neueren Version des Control-Accessories können Sie ebenfalls die RS232 Konfiguration vornehmen. Klicken Sie dazu den Menueintrag "Kontrollfeld" und im darauf erscheinenden Fenster unten links das Feld "RS232-Parameter" an. Es erscheint das Dialogfeld wie in Abb. 14.2.

RS232 KONFIGURATION

Baud Rate:

Parität: Keine Odd Even

Duplex: Voll Halb

Bits/Zchn: 8 7 6 5

Bit 8: An Aus

Kommunikationsparameter

Xon/Xoff: An Aus

Rts/Cts: An Aus

Abb. 14.2. Zur Einstellung der RS232 Schnittstelle

Mit dem VT52 Emulator können Sie vorgängig von Hand testen, ob die Verbindung funk-

tioniert. Für eine detaillierte Beschreibung der einzelnen Parameter lesen Sie bitte in Ihrem Atari-Handbuch unter Kapitel 4.1.2. "VT52 Emulator" und Kapitel 4.1.4. bzw. 4.1.3.6. "RS232 Einstellung" nach.

Hinweis: Das Desktop bietet Ihnen die Möglichkeit die gemachten Veränderungen an der RS232 Konfiguration zu speichern (Sie müssen dazu AUGUR verlassen). Klicken Sie auf den Menueintrag "Arbeit sichern" unter dem Menu "Extras" (im Desktop). Der Atari speichert darauf neben weiteren Optionen alle Einstellungen der RS232 Schnittstelle auf eine Datei mit Namen "DESKTOP.INF" auf die Diskette in Laufwerk A oder, falls Sie mit einer Harddisk arbeiten auf die Partition "C". Befindet sich diese Datei und das Accessory "EMULATOR.ACC" oder die neuere Version von "CONTROL.ACC" auf ihrer Bootdiskette, so wird die RS232 Schnittstelle beim Systemstart automatisch richtig eingestellt.

Wenn Sie mit einem IBM oder kompatiblen Computer mit 3 1/2-Zoll Disketten-Laufwerk arbeiten, oder falls Sie an Ihrem Atari ein 5

1/4-Zoll Laufwerk angeschlossen haben, gibt es noch eine Alternative für den Datentransfer: Formatieren Sie eine Diskette auf dem IBM Computer. Der Atari ist fähig, eine solche Diskette zu lesen oder zu beschreiben. Sie können also mit AUGUR Texte darauf speichern und dann mit dem IBM Computer wieder lesen.

14.5. Verschiedene Textformate

Im Normalfall werden die Texte im ASCII-Format abgespeichert. Dies ist ein weit verbreitetes, standardisiertes Textformat (American Standard Code for Information Interchange). Die allermeisten Textverarbeitungs- und Desktop-Publishing- Programme sind fähig dieses Format zu lesen. Für interessierte Leser ist der entsprechende Code im Anhang E abgebildet.

Für das weitverbreitete Textverarbeitungsprogramm "1st Word Plus" wurde die Möglichkeit geschaffen, Texte in seinem eigenen Format abzuspeichern. Dieses Programm kann zwar Texte im ASCII-Format ebenfalls lesen, die Neuformatierung bereitet jedoch einige Mühe (Carriage>Returns werden als Paragrafenden

betrachtet).

Am Schluss des Menus "Text" befinden sich die beiden Einträge "ASCII" und "1st Word Plus". Mit diesen Einträgen kann gewählt werden, unter welchem der beiden Formate der Text abgespeichert wird. Das aktuelle Format ist im Menu mit einem Haken versehen. Es kann durch Anklicken des entsprechenden Menueintrages ("ASCII" oder "1st Word Plus") gewechselt werden. Beim Starten von AUGUR ist das ASCII-Format aktiv.

Beim Abspeichern eines Textes im 1st Word Plus Format wird auch der Unterschied zwischen Bindestrichen und Trennstrichen berücksichtigt. 1st Word Plus setzt zur Unterscheidung anstelle von Trennstrichen eine Tilde. Solche Tilden können beim Umformatieren entfernt und mit der Trennhilfe neu gesetzt werden. Bindestriche hingegen bleiben im Text immer erhalten. Nun werden in den meisten Schriftarten für Trenn- und Bindestriche die selben Zeichen verwendet und AUGUR kann sie nicht unterscheiden. Die Folge wäre, dass beim Umformatieren des Textes in 1st Word Plus die Trennstriche erhalten blieben, was natürlich unerwünscht ist. AUGUR unterscheidet deshalb Trenn- und Bindestriche automatisch nach folgender Re-

gel: Trennstriche treten immer am Zeilenende auf und zwar nur dann, wenn die nächste Zeile mit einem kleinen Buchstaben beginnt.

Eine weitere Formateigenschaft wird mit der Option "**kein Einrücken**" im Menu "Text" eingestellt. Im Normalfall versucht AUGUR zwischen den Wörtern und am linken Rand so viele Leerzeichen einzufügen wie im Original vorhanden waren. Damit bleiben Tabellen mehr oder weniger gut erhalten. Für viele Anwendungen ist es jedoch besser, wenn zwischen zwei Wörtern nur gerade ein Leerschlag ist und wenn am linken Rand überhaupt keine Leerschläge sind. Es ist dann meistens einfacher, Tabulatoren zu setzen.

Genau dies wird mit "kein Einrücken" bewirkt. Wird der entsprechende Menueintrag angeklickt (er wird darauf mit einem Haken versehen), so werden am linken Rand gar keine und zwischen zwei Wörtern nur ein Leerschlag eingefügt. Durch erneutes Anwählen dieses Menüpunktes kann der Modus wieder ausgeschaltete werden.

Alle Formateinstellungen (ASCII oder 1st Word Plus und "kein Einrücken" gelten für das Anzeigen, Abspeichern, sowie für das Senden über die RS232 Schnittstelle. Beim Anhängen

des erkannten Textes an eine bestehende Textdatei wird automatisch deren Textformat (ASCII oder 1st Word Plus) gewählt. Für diesen Fall hat also die Einstellung ASCII oder 1st Word Plus keine Bedeutung. "Kein Einrücken" dagegen wird auch beim Anhängen an eine Textdatei berücksichtigt.

14.6. Den Text löschen

Gemeint ist das Löschen eines Textes aus dem Arbeitsspeicher des Atari. Wie Sie gegebenenfalls eine erstellte Textdatei auf Diskette wieder löschen können, lesen Sie bitte in Kapitel 15 "Weitere Funktionen" nach.

Der Text im Arbeitsspeicher kann mit dem Menueintrag "löschen" im Menu "Text" gelöscht werden.

Dies ist beispielsweise dann sinnvoll, wenn ein Text erkannt wurde und Sie sich entscheiden, es mit einem anderen Erkennungsmodus noch einmal zu versuchen.

Beim Laden einer Bibliothek wird der Text im Arbeitsspeicher gelöscht, allerdings nicht ohne dass Sie zuvor gewarnt würden. Dies ist not-

wendig, weil der Text in codierten Form im Computer abgespeichert ist. In der Bibliothek befindet sich Information, welche zur Umwandlung des Textes in ASCII-Code benötigt wird. Ohne zugehörige Bibliothek ist der Text also wertlos.

Wollen Sie den aktuellen Text nicht verlieren, so müssen Sie ihn auf Diskette abspeichern.

15. Weitere Funktionen

In diesem Kapitel werden jene Funktionen beschrieben, welche mit der Schrifterkennung direkt nichts zu tun haben, bis jetzt also unerwähnt geblieben sind.

15.1. Das Menu "Desk"

Das Menu Desk nimmt eine Sonderstellung ein. Nur der oberste Eintrag ("AUGUR...", zeigt eine Box mit Informationen über das Programm und den Distributor an) stammt nämlich von AUGUR selbst.

Bei den restlichen Einträgen (unter Umständen sind auch keine weiteren Einträge vorhanden) handelt es sich um selbständige Programme, welche beim Aufstarten des Systems geladen worden sind und resident im Arbeitsspeicher bleiben, bis der Computer wieder ausgeschaltet wird. Diese Programme werden **Desk-Accessories** genannt (auf Deutsch: Schreibtischzubehöre). Sie können über die entsprechenden Menueinträge aufgerufen werden.

Es handelt sich meistens um Hilfsprogramme,

wie zum Beispiel eine Uhr oder einen Taschenrechner. Das praktische daran ist, dass Sie gestartet werden können, ohne dass AUGUR verlassen werden muss. Sie laufen also quasi parallel zu AUGUR. Sie können auch von den meisten anderen Programmen aus aufgerufen werden.

Accessories tragen den Extender ".ACC". Jene Accessories, welche sich beim Einschalten des Computers auf der Diskette im Laufwerk A (Boot Diskette) befinden werden automatisch geladen. Sie können sich also das Menu "Desk" selbst zusammenstellen.

Auf der mitgelieferten Beispiel-Diskette, sowie auf der zum Lieferumfang des Atari gehörenden Language Disk, befinden sich die beiden Accessories "CONTROL.ACC" und "EMULATOR".ACC. Sie bewirken die vier Menueinträge "Kontrollfeld", "Drucker Anpassung", "VT52 Emulator" und "RS 232 Einstellung" (bei der neuen Version dieser Programme nur noch "Kontrollfeld" und "VT52 Emulator". Die genaue Funktionsweise dieser Accessories ist in Ihrem Atarihandbuch im Kapitel 4.1. Das Menu "Desk" beschrieben.

AUGUR macht direkten gebrauch vom Emulator- Accessory (siehe Kapitel 14.4). Zu er-

wähnen ist noch das CP14LIST-Accessory, welches mit dem Hawk Scanner geliefert wird. Mit seiner Hilfe ist es möglich, von AUGUR erkannte und auf Diskette gespeicherte Texte sehr schnell auf dem eingebauten Thermo-drucker des Hawk auszudrucken.

15.2. Eine Datei löschen

Dieser Befehl befindet sich im Menu "Datei". Er erlaubt es, Dateien auf Diskette oder Hard-disk zu löschen. Vorsicht: eine gelöschte Datei ist für immer verloren! Vor allem Bilder benötigen sehr viel Speicherplatz. Sie sind nach dem Erkennen auch meistens nicht mehr notwendig. Ein Bild löschen schafft Platz für viele Texte und Bibliotheken.

Wenn der Befehl "Datei löschen" angeklickt wird, erscheint die Objektauswahl-Box, in welcher die zu löschende Datei ausgewählt werden kann. Gelöscht wird die Datei aber erst nach einer weiteren Sicherheitsabfrage.

15.3. Eine Diskette formatieren

Bevor Dateien auf eine neue Diskette kopiert werden können, muss diese formatiert werden. Bei diesem Vorgang werden Daten auf die Diskette geschrieben, welche später für das Speichern und Laden von Dateien gebraucht werden.

Legen Sie die zu formatierende Diskette in Laufwerk A. Wählen Sie danach mit der Maus den Menüpunkt "**Diskette A formatieren**" im Menu "Datei" an. Sie können auch eine bereits beschriebene Diskette formatieren. Dabei werden jedoch alle darauf enthaltenen Dateien zerstört. Deshalb startet der Formatiervorgang erst nach einer zusätzlichen Sicherheitsabfrage.

Während des Formatiervorganges erscheint unter dem Titel "Arbeit" (links auf dem Bildschirm) das Wort "Formatieren:" und dahinter eine Zahl, welche von 80 an rückwärts bis auf 0 zählt. Es handelt sich dabei um die Spurnummer. Bei 0 ist der Formatiervorgang beendet.

Hinweis: AUGUR kann Disketten nur zweiseitig formatieren. Wollen Sie Disketten einseitig formatieren, so müssen Sie

dies vom Desktop aus tun. Die zugehörige Beschreibung finden Sie im Atari-Handbuch, Kapitel 4.2.5. "formatiere". Dies ist allerdings nur dann nötig, wenn Sie einseitige Disketten kaufen oder ein einseitiges Diskettenlaufwerk verwenden (von den original Atari Laufwerken nur das SF354. Eingebaute Laufwerke sind alle zweiseitig).

15.4. Serierarbeit

Die Funktion Serierarbeit dient der einfacheren Verarbeitung von Dokumenten, welche viele Seiten lang sind. AUGUR übernimmt dabei alle nötigen Schritte für die Schrifterkennung. Die Aufgabe des Benützers ist nur noch die Seiten unter dem Scanner zu wechseln, während AUGUR am Erkennen ist. Treten unbekannte Buchstaben auf, so müssen diese am Ende jeder Seite eingetippt werden.

Nach dem Anklicken des Eintrages "**Serierarbeit**" im Menu "Schrifterkennung" werden vollautomatisch folgende Funktionen nacheinander aufgerufen: "Bild scannen", "Text erkennen", "Schrift lernen" (nur wenn neue

Buchstaben in die Bibliothek aufgenommen wurden), "Text anhängen". Danach wird ohne Halt wieder bei "Bild scannen" weitergefahren. Bis dahin sollte der Benutzer die Vorlage also gewechselt haben.

Ist die Option "Linien immer entfernen" (siehe unten) aktiviert, so werden die wagrechten und senkrechten Linien immer vor dem Erkennen entfernt. Die Option "Rahmen erhalten" wird bei der Seriearbeit immer aktiviert.

Durch einen Tastendruck der Tastatur, während AUGUR erkennt, kann die Seriearbeit abgebrochen werden. Eine Alternative dazu ist es, einfach kein neues Dokument unter den Scanner zu legen. Beim nächsten Scenvorgang wird entweder ein ganz schwarzes oder ganz weisses Bild gescannt. In beiden Fällen wird AUGUR reklamieren: "keine Zeilen gefunden" und die Möglichkeit offen lassen, die Seriearbeit abzuberechnen.

Bevor die Seriearbeit jedoch gestartet werden kann, muss AUGUR noch mitgeteilt werden, an welche Textdatei der erkannte Text jeweils angehängt werden soll. Dazu muss die erste Seite des Dokumentes mit "Text erkennen" erkannt werden und mit "Text anhängen" an eine bestehende Datei angehängt werden oder

mit "Text sichern" eine neue Textdatei erstellt werden. Beim Starten der Seriarbeit für die folgenden Seiten des Dokumentes wird der erkannte Text immer an diese Datei angehängt.

15.5. Fixed Space

Ist der Fixed Space Modus aktiv (Eintrag "**Fixed Space**" im Menu "Schrifterkennung"), so wird AUGUR gezwungen, alle Buchstabenabstände als gleich zu betrachten. Nicht zusammenhängende Zeichen, wie zum Beispiel Anführungs- und Schlusszeichen werden dann wieder zusammengefügt. Dies funktioniert natürlich nur, wenn die Schrift wirklich konstante Buchstabenabstände hat (Fixed Spaced Schrift), also zum Beispiel bei Schreibmaschinenschrift. Im Normalfall unterscheidet AUGUR jedoch selbst zwischen Fixed Spaced- und Proportionalschrift. Dieser Modus muss also nur angewandt werden, wenn zu viele Zeichen nicht zusammenhängen (sehr viele Anführungszeichen, schlechte Kopie). AUGUR könnte sonst, bedingt durch die sehr unterschiedlichen Zeichenabstände, die Schrift für proportional halten.

Der Fixed Space Modus wird durch Anklicken des Menueintrages "**Fixed Space**" im Menu "Diverses" ein und ausgeschaltet. Ist er aktiv, so wird der entsprechende Menueintrag mit einem Haken versehen.

15.6. Linien entfernen

Diese Funktion entfernt senkrechte und wagrechte gerade Linien. Ausgelöst wird sie durch Anklicken von "**Linien entfernen**" im Menu "Diverses". Sie wirkt sich, unabhängig von den eingezeichneten Rahmen, immer auf das ganze Bild aus.

Damit ist es möglich, Tabellen zu lesen, welche von vertikalen Strichen durchzogen sind. Oft ist es auch nützlich vertikale Striche, welche Zeitungsspalten separieren zu entfernen. Die Rahmen lassen sich danach meist einfacher einzeichnen.

Sie kann ebenfalls dazu verwendet werden, Unterstreichestrache zu entfernen. Unterstreichestrache sollten immer dann vor dem Erkennen des Dokumentes entfernt werden, wenn sie den Text berühren. AUGUR könnte dann nämlich die Buchstaben nicht mehr sicher tren-

nen.

Nach dem Anklicken von "Linien entfernen" erscheint eine Box, welche nach der Buchstaben-grösse fragt (klein, mittel oder gross). Diese Angabe benötigt AUGUR um zu entscheiden, wie lang eine Linie mindestens sein muss, damit sie sicher nicht Teil eines Buchstabens ist. Linien, welche länger als diese Grenze sind, werden entfernt. Diese Grenze ist natürlich auch von der Auflösung des Scanners abhängig. Bei 200 Dpi entsprechen die Angaben klein, mittel und gross ungefähr 5, 10 und 20 mm. Bei 400 Dpi sind dies 2.5, 5 und 10 mm. Wird bei einer grossen Schrift (oder grossen Auflösung des Scanners) ein zu kleiner Wert angegeben, so können die Buchstaben "angefressen" werden.

Entfernt werden nur senkrechte und wagrechte gerade Linien. Schräge und krumme Linien bleiben erhalten.

Auf Wunsch kann diese Funktion automatisch vor jedem Erkennen eines Dokumentes ausgelöst werden. Klicken Sie dazu auf den Menüeintrag "**Linien immer entf.**" im Menü "Diverses". Dieser Eintrag wird darauf mit einem Haken versehen und die Funktion "Linien Entfernen" wird ab jetzt, bei jedem Aufruf der Schrifterkennung, automatisch aus-

gelöst. Durch erneutes Anklicken von "Linien immer entf." wird diese Automatik wieder ausgeschaltet.

Beim automatischen Linienentfernen wird nicht mehr nach der Buchstabengröße gefragt. Es gilt der, mit "Linien entfernen" zuletzt eingestellte Wert.

15.7. Kursiv aufrichten

Falls der ganze Text in einem Rahmen kursiv geschrieben ist, kann er vor der eigentlichen Erkennung mit dieser Option aufgerichtet werden (Menu "Diverses"). Dadurch wird die Separierung der einzelnen Zeichen etwas beschleunigt und verbessert. Es führt jedoch zu keiner Qualitätseinbusse, wenn diese Option weggelassen wird.

15.8. Das Programm verlassen

AUGUR kann mit dem Menueintrag "**Programm verlassen**" im Menu "Datei" verlassen werden. Ist eine Bibliothek erweitert oder neu erstellt und noch nicht gespeichert

oder ist der erkannte Text noch nicht gespeichert worden, so erscheint vor dem Verlassen eine Warnbox, welche die Möglichkeiten offen lässt, die entsprechende Datei zu speichern oder noch einmal ins Programm zurückzukehren.

16. Tips zum Gebrauch von Textverarbeitungsprogrammen

Die meisten Textverarbeitungsprogramme bieten die Möglichkeit einer Wortkorrektur. Davon sollte auf jeden Fall Gebrauch gemacht werden, um von AUGUR eventuell begangene Fehler leicht und sicher zu finden.

Es gibt Schriftarten, bei denen die Buchstaben gross "I" und klein "l" sehr ähnlich oder sogar gleich sind und AUGUR sie nicht unterscheiden kann. In diesem Fall sollte im Lernmodus klein "l" eingegeben werden. Da grosse "I" nur an einem Wortanfang stehen können, kann im Textverarbeitungsprogramm mit der Funktion "Suchen und Ersetzen" die Kombination " l" (Leerschlag + "l") wo nötig gegen " I" (Leerschlag + "I") ausgetauscht werden.

Weitere Schwierigkeiten kann das Anführungszeichen (") bereiten. Bei Proportionalschrift wird dies AUGUR bestimmt für zwei Hochkommas halten ('). Bei Fixed Spaced Schriften (eventuell auf Fixed Space gezwungen) sollte dieses Problem hingegen nicht auftreten. Hier hat AUGUR die Möglichkeit, zu erkennen,

dass es sich in Wirklichkeit um ein einzelnes Zeichen handelt. Abhilfe schafft auch hier die Funktion "Suchen und Ersetzen" ihrer Textverarbeitung. Suchen Sie nach zwei unmittelbar hintereinanderliegenden Hochkommas (") und ersetzen Sie diese durch das Anführungszeichen (").

1.1 Blocksatz gedruckte Vorlagen haben zwischen den Wörtern zum Teil mehrere Leerschläge. Es kann also sein, dass AUGUR zwei Wörter mit mehr als einem Leerschlag trennt. Ist dies unerwünscht, so hilft folgendes Vorgehen: Verwenden Sie die Funktion "Suchen und Ersetzen" in Ihrem Textverarbeitungsprogramm und ersetzen Sie "" (zwei Leerschläge) gegen " " (ein Leerschlag). Wiederholen sie diesen Schritt, bis keine doppelten Leerschläge mehr vorkommen.

Anhang A:

Zusammenfassung der Befehle in der Menuleiste

1. Das Menu "Desk"

Augur... Eine Box mit Informationen über das Programm und den Distributor wird angezeigt.

Accessories Unterhalb von "Augur..." befinden sich die sogenannten Desk Accessories (auf Deutsch: Schreibtisch-zubehör). Dies sind Hilfsprogramme, welche parallel zu AUGUR im Speicher stehen und mit diesen Menüpunkten aufgerufen werden können. Die Anzahl ist variabel.

2. Das Menu "Datei"

Bild scannen Damit wird ein Dokument abgetastet (gescannt). Beim Hawk CP14 muss danach der Startknopf (rosa Knopf

auf dem Bedienfeld des Scanners) gedrückt werden, der Hawk 432 startet automatisch.

Bild laden Ein Bild kann von einem Massenspeicher (Diskette, Harddisk, ...) in das Programm geladen werden.

— **Bild sichern** Das aktuelle Bild wird auf einen Massenspeicher geschrieben.

Bibliothek laden Eine neue Bibliothek wird geladen. In einer Bibliothek befinden sich Prototypen von den Buchstaben der jeweiligen Schriftart.

Bibliothek sichern Die aktuelle Bibliothek wird auf einen Massenspeicher gespeichert.

— **Bibliothek löschen** Die aktuelle Bibliothek wird gelöscht. Dies ist dann sinnvoll, wenn ein neues Dokument gelesen werden soll, dessen Schriftart verschieden vom Inhalt der aktuellen Bibliothek ist. Die Bibliothek würde dann nichts zum Erkennungsvorgang beitragen, sondern diesen nur verlangsamen.

Datei löschen Eine Datei wird von einem

Massenspeicher gelöscht.

Disk A formatieren Die Diskette im Laufwerk A wird zweiseitig formatiert. Dieser Vorgang muss mit allen neuen Disketten einmal gemacht werden, bevor sie mit Daten beschrieben werden können.

Programm verlassen Das Programm wird verlassen. Ist ein erkannter Text oder eine neu erstellte oder veränderte Bibliothek noch nicht abgespeichert worden, so erscheint vor dem endgültigen Verlassen eine Warnbox.

3. Das Menu "Text"

sichern Der erkannte Text wird mit dem eingestellten Format (ASCII oder 1st Word Plus) auf einen Massenspeicher geschrieben.

anhängen Der erkannte Text wird einer bereits existierenden Textdatei angehängt. Als Format (ASCII oder 1st Word Plus) wird dabei jenes der bereits existierenden Textdatei genommen, unabhängig davon, was im Programm eingestellt ist.

senden Der erkannte Text wird mit dem eingestellten Format (ASCII oder 1st Word Plus) über die RS232- Schnittstelle (Modem-Port) an einen anderen Computer gesendet. Die RS232-Konfiguration kann mit dem Accessory "EMULATOR.ACC" oder mit einer neueren Version des Accessories "CONTROL.ACC" eingestellt werden. Diese Option dient dazu, dass auf einem anderen Computer (z.B. Macintosh) bereits am Text gearbeitet werden kann, während AUGUR die nächsten Seiten erkennt.

anzeigen Der erkannte Text wird auf dem Bildschirm angezeigt. Hat nicht der gesamte Text auf einmal Platz, so kann mit einem Tastendruck (Maus oder Tastatur) zur nächsten Bildschirmseite vorgerückt werden. Mit [Esc] oder einem Tastendruck nach der letzten Bildschirmseite wird dieser Modus wieder verlassen.

löschen Der erkannte Text wird gelöscht. Damit wird der Textpuffer frei für neuen Text.

ASCII Bei "Text sichern" wird das ASCII-Format verwendet.

1st Word Plus Bei "Text sichern" wird das 1st Word Plus-Format verwendet. Fast alle Textverarbeitungs- und Desktop-Publishing-Programme können ASCII-Dateien lesen. Da das weitverbreitete Programm "1st Word Plus" jedoch Mühe hat, solche Dateien neu zu formatieren, wurde die Möglichkeit geschaffen, Texte in seinem eigenen Format zu speichern.

kein Einrücken Wird diese Funktion angeklickt (sie wird darauf mit einem Haken versehen), so werden am linken Rand gar keine und zwischen den Wörtern nur gerade ein Leerzeichen eingefügt. Dies ist meistens dann sinnvoll, wenn der Text mit einem Textverarbeitungsprogramm neu formatiert und eigene Tabulatoren gesetzt werden sollen.

Durch erneutes Anklicken dieses Menüeintrages wird die Funktion wieder ausgeschaltet.

Rahmenmarker Wird diese Funktion angeklickt (sie wird darauf mit einem Haken versehen), so wird beim Speichern des Textes oder Senden über die RS232 Schnittstelle zwischen den Text aus verschiedenen Rahmen jeweils ein Formfeed-

Zeichen gesetzt (12 oder in Hex C). Dieses Zeichen kann bei der Aufbereitung erkannter Texte für Datenbanken von Nutzen sein. Die Funktion gilt sowohl für das ASCII, als auch für das 1st Word Plus Format.

Durch erneutes Anklicken dieses Menüeintrages wird die Funktion wieder ausgeschaltet.

4. Das Menu "Rahmen"

Rahmen bezeichnen jene Stellen im Bild, welche erkannt werden sollen. Sie werden mit der Maus im verkleinerten Bild eingezeichnet (linke Maustaste gedrückt halten). Klicken Sie nur kurz in das verkleinerte Bild, so erhalten Sie eine Vergrößerung über den ganzen Bildschirm (Scroll-Modus). Mit der Bewegung der Maus verschiebt sich auch der angezeigte Bildausschnitt. Mit den Cursorpfeilen (auf der Tastatur) kann jetzt die linke obere Rahmenecke einjustiert werden. Mit "Shift" und den Cursorpfeilen, wird die rechte untere Ecke korrigiert. Dieser Modus kann mit einem Tastendruck (Maus oder Tastatur) wieder verlassen werden.

Es sind mehrere Rahmen möglich. Sie werden in jener Reihenfolge erkannt, in welcher sie definiert worden sind. Ist kein Rahmen eingezeichnet, so wird das gesamte Bild erkannt.

Rahmen löschen Der zuletzt gezeichnete Rahmen wird gelöscht. Diese Funktion kann auch mit der Taste "Esc" ausgelöst werden.

alle Rahmen löschen Alle Rahmen werden gelöscht.

Rahmen erhalten Die aufgespannten Rahmen werden beim Laden oder Scannen eines neuen Bildes nicht gelöscht. Dies ist zum Beispiel bei mehreren Formularen mit immer gleichen Spaltenanordnungen sinnvoll.

Wissenswertes... Ein Dialogfeld mit Informationen über Rahmen wird gezeigt.

5. Das Menu "Schrifterkennung"

Text erkennen Das geladene oder gescannte Bild wird erkannt.

Schrift lernen Jene Buchstaben, welche in der Bibliothek noch nicht vorhanden sind, werden gezeigt und können vom Benutzer über die Tastatur eingegeben werden.

Text anzeigen Der erkannte Text wird auf dem Bildschirm angezeigt. Es handelt sich um die gleiche Funktion wie unter dem Menutitel "Text".

Seriearbeit damit wird der automatische Erkennungsablauf gestartet. Dabei werden vollautomatisch folgende Funktionen nacheinander aufgerufen: "Bild scannen", "Text erkennen", "Schrift lernen" (nur wenn neue Buchstaben in die Bibliothek aufgenommen wurden), "Text anhängen". Danach wird ohne Halt wieder bei "Bild scannen" weitergefahren. Bis dahin sollte der Benutzer die Vorlage also gewechselt haben.

Ist die Option "Linien immer entfernen" aktiviert, so werden die wagrechten und senkrechten Linien immer vor dem Erkennen entfernt. Die Option "Rahmen erhalten" wird bei der Seriearbeit immer aktiviert.

Die Seriearbeit bringt vor allem dann Nutzen, wenn viele Blätter mit der glei-

chen Schriftart nacheinander erkannt werden sollen.

Durch einen Tastendruck während AUGUR erkennt, kann die Seriearbeit abgebrochen werden.

schneller Modus AUGUR liest in diesem Modus schnell, jedoch nicht so präzise. Er ist vor allem für grosse Auflösungen (400 Dpi) geeignet. Bei kleineren Auflösungen kann er bei grossen Schriften oder bei klassischer Schreibmaschinenschrift (Courier) verwendet werden. Bei etwas "verzitterten" Schriften, wie zum Beispiel Matrixdruckerschrift, führt dieser Modus meistens zu besseren Resultaten, als der präzise Modus (siehe "präziser Modus").

goldenes Mittel Dieser Modus kann für moderne und kleine Schreibmaschinenschriften bei mittlerer und kleiner Scannerauflösung (300, 200 Dpi) verwendet werden. Ebenso ist er brauchbar für gross geschriebene Drucksachen.

präziser Modus In diesem Modus arbeitet AUGUR sehr genau und ist dementsprechend etwas langsamer. Er ist bei klein geschriebenen Drucksachen (wie

z.B. Zeitung) und kleiner Auflösung des Scanners (200 Dpi) angebracht. Vorsichtshalber kann bei 200 Dpi immer mit diesem Modus begonnen werden. Bei unscharfen oder verzitterten Schriften, sowie bei ungleichmässig belichteten Kopien kann es jedoch vorkommen, dass AUGUR gleiche Buchstaben für verschieden hält und die Bibliothek stark anwächst. In diesem Fall sollte auf einen schnelleren Modus umgeschaltet werden.

Fixed Space Ist "Fixed Space" aktiv, so wird AUGUR gezwungen, alle Buchstabenabstände als gleich zu betrachten. Nicht zusammenhängende Zeichen, wie zum Beispiel Anführungs- und Schlusszeichen werden dann wieder zusammengefügt. Im Normalfall unterscheidet AUGUR jedoch selbst zwischen Fixed Spaced- und Proportionalschrift. Dieser Modus muss also nur angewandt werden, wenn zu viele Zeichen nicht zusammenhängen (sehr viele Anführungszeichen, schlechte Kopie). AUGUR könnte sonst, bedingt durch die sehr unterschiedlichen Zeichenabstände, die Schrift für proportional halten.

6. Das Menu "Diverses"

Linien entfernen Diese Funktion entfernt auf dem ganzen Bild, unabhängig von den eingezeichneten Rahmen, wagrechte und senkrechte gerade Linien. Dies ist nützlich, um störende Unterstreichestrüche oder senkrechte Linien, welche Spalten oder Tabellen separieren, zu entfernen.

Linien immer entf. Wird dieser Menueintrag angeklickt (er wird darauf mit einem Haken versehen), so wird die obige Funktion "Linien entfernen" automatisch bei jedem Erkennen des Textes ausgeführt.

Durch erneutes Anklicken des Menueintrages wird die Funktion wieder ausgeschaltet.

Kursiv aufrichten Falls der Text in einem Rahmen vollständig kursiv geschrieben ist, kann er vor der eigentlichen Erkennung mit dieser Option aufgerichtet werden. Dadurch wird die Separierung der einzelnen Zeichen etwas beschleunigt und verbessert. Es führt jedoch zu keiner

Qualitätseinbusse, wenn diese Option weggelassen wird.

Hochformat Vorlagen werden mit dem Kopf nach rechts unter den Scanner gelegt.

Querformat Vorlagen werden mit dem Kopf nach hinten unter den Scanner gelegt, also quer. Dies kann bei Büchern nützlich sein. Es können zwei Buchseiten auf einmal bearbeitet werden.

Scanner Wird diese Funktion angeklickt, so erscheint ein Dialogfeld (Abb. 6.1) in welcher die Einstellungen (Helligkeit, Auflösung) für den Hawk 432 Scanner vorgenommen werden können. Es kann auch zwischen den Scannern Hawk CP14 und Hawk 432 umgeschaltete werden (nur sinnvoll, wenn beide Scanner gleichzeitig angeschlossen sind). Beim Hawk CP124 muss die Einstellung der Helligkeit am Scanner selbst vorgenommen werden.

Anhang B:

Zusammenfassung der Befehle im Lernmodus

Im Lernmodus ist es möglich, den erkannten Zeichen richtige Buchstaben zuzuweisen (AUGUR lernt sie dabei) oder einfach nur den Inhalt einer Bibliothek sichtbar zu machen.

Der Lernmodus wird durch den Menueintrag "Schrift lernen" im Menu "Schrifterkennung" aufgerufen.

[Esc]: Mit dieser Taste wird der Lernmodus wieder verlassen.

[Return] oder **[Enter]:** Damit ist es möglich, verwachsenen Zeichen mehrere Buchstaben zuzuordnen (maximal 4). Abgeschlossen wird die Zeicheneingabe wieder mit [Return] oder [Enter].

[Del]: Dem entsprechenden Zeichen wird kein Buchstabe zugewiesen. Im Text wird an diesen Stellen nichts eingefügt, auch

kein Leerschlag.

[Back]: Springt zum letzten Buchstaben zurück. Bei der Eingabe einer Zeichenkette (siehe [Return]) lassen sich mit [Back] der Reihe nach die eingegebenen Buchstaben wieder löschen.

[Cursor links]: Springt zum letzten Buchstaben zurück. Wird der linke Cursorpfeil während der Eingabe einer Zeichenkette (siehe [Return]) gedrückt, so wird ebenfalls ein Zeichen zurückgesprungen, die gemachten Änderungen werden dabei allerdings rückgängig gemacht.

[Cursor rechts]: Springt zum nächsten Buchstaben. Wird der rechte Cursorpfeil während der Eingabe einer Zeichenkette (siehe [Return]) gedrückt, so wird ebenfalls zum nächsten Zeichen gesprungen, die gemachten Änderungen werden dabei allerdings rückgängig gemacht.

[Shift Cursors]: Wird die [Shift]-Taste und gleichzeitig der linke oder der rechte Cursorpfeil gedrückt, so wird jeweils 10 Zeichen zurück oder vorwärts gesprungen.

[Home]: Springt zum ersten Zeichen in der Bibliothek.

[F1] - [F6]: Dies sind sogenannte Totzeichen. Das entsprechende Menu befindet sich auf der untersten Zeile des linken Fensters auf dem Bildschirm. Soll beispielsweise ein "é" eingegeben werden, so drücken Sie zuerst [F1] und anschliessend das "e".

Sonderzeichen: Sie sind im linken Fenster auf dem Bildschirm zusammengestellt und werden mit Klicken der Maus eingegeben. Eine Alternative dazu sind die Funktionstasten F1 - F6 (siehe oben).

Anhang C: Alphabetische Begriffserklärungen

Begriffe, welche selbst in diesem Anhang erscheinen, sind in den Erklärungen fett gedruckt.

abtasten: (scannen) Beim Abtasten wird eine für den Computer lesbare "Kopie" der Vorlage erstellt. Dazu wird das Original in einzelne Bildpunkte zerlegt. In unserem Fall kann jeder Bildpunkt nur noch die Farbe Schwarz oder Weiss haben.

Accessory: So werden spezielle Hilfsprogramme genannt. Sie tragen den **Extender ".ACC"** und werden beim Einschalten des Computers automatisch von der **Boot-Diskette** geladen und gestartet. Sie laufen parallel zu anderen Anwendungen und können von den meisten Programmen unter dem Menutitel "Desk" aufgerufen werden. Beispiele dafür sind die Accessories "CONTROL.ACC" und "EMULATOR-

.ACC", welche auf der Language Disk Ihres Ataris mitgeliefert werden (Siehe auch Kapitel 4.1 "Das Menü Desk" im Atari-Handbuch).

Bibliothek: Die Bibliothek dient zur Erkennung eines Textes. In ihr befinden sich Prototypen von den Buchstaben der jeweiligen Schriftart oder von verschiedenen Schriftarten. Abhängig von der Buchstabengröße kann eine Bibliothek maximal ungefähr 1000 verschiedene Buchstaben enthalten. Eine Bibliothek kann sich im Programmspeicher von AUGUR oder auf einem **Massenspeicher** befinden.

Boot-Diskette: Als Boot-Diskette wird jene Diskette bezeichnet, welche sich beim Einschalten des Computers im Laufwerk A (eingebautes Laufwerk) befindet. Von ihr werden die **Accessories**, sowie die Programme, welche sich im Ordner namens "AUTO" befinden automatisch geladen und gestartet. Damit ist es möglich, das System auf gewisse Ansprüche individuell anzupassen.

Cartridge: Neben der **Diskette** eine weitere Möglichkeit, um Programme zu gebrauchen. Cartridges sind eine Art Stecker,

die auf der linken Computerseite eingesteckt werden. Solche Programme arbeiten extern, brauchen also nicht erst geladen zu werden. Zudem wird für den Programmcode kein zusätzlicher Arbeitsspeicher benötigt.

Datei: Eine Zusammenfassung von Informationen, welche auf einem **Massenspeicher** abgelegt sind. Dateien werden vom **Desktop** als Bilder oder als Text angezeigt. AUGUR verwendet Dateien zur Speicherung der Bilder, Bibliotheken und der erkannten Texte.

Desk Accessory: Siehe **Accessory**

Desktop: (Schreibtischoberfläche) So wird jenes Programm genannt, welches normalerweise nach dem Einschalten des Computers erscheint. Es handelt sich um eine grafische Benutzeroberfläche. **Disketten-Laufwerke** und **Dateien** sind als Bilder dargestellt. Vom Desktop aus können weitere Programme gestartet, Dateien kopiert oder gelöscht werden.

Desktop-Publishing: Darunter versteht man die computergestützte Erzeugung von Drucksachen, wie Zeitungen, Bücher oder

Prospekte. Dabei wird das Dokument auf dem Bildschirm so angezeigt, wie es später ausgedruckt wird.

Diskette: Eine magnetisch beschichtete Scheibe, welche von einem Kunststoffumschlag geschützt wird. Sie dient zur Datenspeicherung. Auf dem Atari werden 3 1/2 Zoll (= 9 cm) grosse Disketten verwendet.

Disketten-Laufwerk: Von hier werden die **Disketten** beschrieben und gelesen. Auf dem Atari 1040ST befindet sich das Disketten-Laufwerk auf der rechten Seite, beim Atari MEGA ST befindet es sich auf der Frontseite rechts. Disketten werden mit dem Metallschieber voran und der Aufschrift nach oben in die Laufwerksöffnung eingeschoben. Sie müssen hörbar einrasten und ganz im Laufwerk verschwinden. Durch Drücken auf den breiten Knopf unterhalb des Laufwerks kann die Diskette wieder ausgeworfen werden. Wichtig: nehmen Sie niemals eine Diskette aus dem Laufwerk, während dieses in Betrieb ist (wird durch ein grünes oder ein rotes Lämpchen angezeigt). Die Information auf der Diskette könnte dabei zerstört werden.

Doppelklick: Bewegen Sie den Mauszeiger (normalerweise ein Pfeil) auf das angegebene Objekt und drücken Sie zweimal sehr kurz hintereinander die linke Maustaste. Die Geschwindigkeit in der dieser Doppelklick ausgeführt werden muss, kann mit dem Kontrollfeld eingestellt werden. Anweisungen darüber erhalten Sie aus Ihrem Atari-Handbuch im Kapitel 4.1.3.3. "Maustasten-Empfindlichkeit".

Dpi: Dots per Inch (= Punkte pro Zoll). Masseneinheit, welche für die Angabe der Auflösung von Scannern und Druckern üblich ist. 200 Dpi = 8 Punkte/mm, 300 Dpi = 12 Punkte/mm und 400 Dpi = 16 Punkte/mm.

Extender: Dies sind drei Buchstaben, welche dem Dateinamen angefügt werden (mit einem Punkt abgetrennt). Sie geben normalerweise Auskunft über die Art der entsprechenden Datei. In AUGUR beispielsweise werden Bilder immer mit dem Extender ".SCN" (= Scanner-Bild), Bibliotheken mit ".LIB" (= Library) und Texte mit ".TXT" (= Text) abgespeichert.

Festplatte: Siehe **Harddisk**

File: Dies ist der Englische Ausdruck für **Datei**.

Fixed Spaces: Bezeichnung für Schriftarten mit konstanten Buchstabenabständen.

Beispiel: "Dies ist die Schreibmaschinenschrift Courier."

Der Gegensatz dazu ist die **Proportionschrift**.

Floppy Disk: Siehe **Diskette**

Floppy Laufwerk: Siehe **Disketten-Laufwerk**.

Font: Schriftart

formatieren: Jede neue Diskette muss vor dem ersten Beschreiben formatiert werden (AUGUR Handbuch: Kapitel 15.3 oder Atari-Handbuch: Kapitel 4.2.5. "formatiere").

Hard Disk: (Festplatte) Fest installierte Speichereinheit auf der Basis einer Magnetplatte, welche als Zusatz an den Computer angeschlossen werden kann. Die Platte selbst kann nicht ausgewechselt werden, wie beim

Disketten-Laufwerk, kann aber erheblich mehr Daten speichern. Zudem ist die Zugriffszeit (Zeit um gewisse Daten zu speichern oder zu lesen) wesentlich kürzer.

Header: Hier ist der Header eines Bildes gemeint. Darin sind Angaben über Grösse und Format des Bildes enthalten. Der Bild-Header befindet sich am Anfang einer Bild-Datei, vor den eigentlichen Bilddaten.

IDC: Intelligent Driver Concept. Stellt einen Geräteunabhängigen Standard für Scanner und Drucker dar. Ein Accessory übernimmt dabei die Steuerung des Ein- oder Ausgabegerätes. Das Anwenderprogramm kann über eine definierte Softwareschnittstelle mit diesem Accessory kommunizieren. Konkret bedeutet dies, dass Sie jeden Scanner, zu dem ein IDC-Treiber existiert, direkt in AUGUR verwenden können.

Eine detaillierte Beschreibung des IDC-Standartes erhalten Sie bei der Marvin AG.

Interface (-box): Damit wird ein elektronisches Gerät bezeichnet, welches die

Anpassung vom Comuter an ein **Peripheriegerät** übernimmt. Ein solches Interface wird zusammen mit dem Hawk 432 geliefert. Der Hawk CP14 benötigt kein Interface.

klicken: Bewegen Sie den Mauszeiger (normalerweise ein Pfeil) auf das angegebene Objekt und drücken Sie danach einmal kurz die linke Maustaste.

Laufwerk A: Dies ist der Name für das eingebaute Diskettenlaufwerk (Der Atari 520+ besitzt kein eingebautes Diskettenlaufwerk, dort ist das Laufwerk A also extern).

Massenspeicher: Damit ist hier eine **Diskette**, **Hard Disk** oder **RAM-Disk** gemeint.

Objektauswahl-Box: Ein Dialogfeld mit dessen Hilfe Namen zum Speichern oder Laden von Dateien eingegeben werden können. Bereits auf Diskette befindliche Dateien werden in alphabetischer Reihenfolge angezeigt und können gegebenenfalls mit der Maus angewählt werden. Die Objektauswahl-Box ist Bestandteil des Atari Betriebssystems und kann demzu-

folge auch von anderen Programmen verwendet werden. AUGUR verwendet sie zum Speichern und Laden von Bildern, Bibliotheken sowie zum Speichern der erkannten Texte (siehe Kapitel 5.4).

Ordner: Zusammenfassung von mehreren Dateien. Ordner bieten die Möglichkeit Dateien auf einer Diskette zu organisieren und zusammenzufassen. Sie werden vom **Desktop** aus erstellt und gelöscht. Ordner-Bilder sehen aus wie Hängeregister-Einlagen. Für mehr Information lesen Sie bitte in Ihrem Atarihandbuch Kapitel 4.2.3. "neuer Ordner" nach.

Peripheriegerät: Ein Zusatzgerät, welches an den Computer angeschlossen werden kann. Beispiele sind Drucker und Scanner.

Pfad: Der Pfadname beinhaltet die Laufwerkskennung (A, B, C...), die **Ordner**namen sowie eventuell ein Dateiname.

Proportionalschrift: Schriftart mit variablen Buchstabenabständen. Dieses Handbuch ist in Proportionalschrift gedruckt. Gegensatz dazu sind **Fixed**

Spaced Schriftarten.

Rahmen: Rahmen bezeichnen jene Stellen im Bild, welche erkannt werden sollen. Sie werden mit der Maus im verkleinerten Bild eingezeichnet (linke Maustaste gedrückt halten). Im **Scroll-Modus** kann mit den Cursorpfeilen die linke obere Rahmenecke feinjustiert werden. Mit [Shift] und den Cursorpfeilen, wird die rechte untere Ecke korrigiert. Es sind mehrere Rahmen möglich. Sie werden in jener Reihenfolge erkannt, in welcher sie eingezeichnet wurden.

RAM: Random Access Memory. Damit ist der interne Arbeitsspeicher des Atari gemeint.

RAM-Disk: Ein sich resident im Speicher des Atari (**RAM**) befindendes Programm, das dem Betriebssystem eine Diskettenstation vortäuscht. Damit können Daten sehr schnell abgespeichert und geladen werden. Sie gehen jedoch beim Ausschalten des Computers verloren und kürzen die Speichergrösse des Systems. RAM-Disks werden von speziellen Hilfsprogrammen installiert.

scannen: Siehe **abtasten**

Schreibtischzubehör: Siehe **Accessory**

Scrollmodus: In diesem Modus wird ein Ausschnitt des geladenen oder gescannten Bildes vergrößert auf dem ganzen Bildschirm dargestellt. Mit der Bewegung der Maus verändert sich auch der angezeigte Bildausschnitt (Scrolling).

In diesen Modus gelangen Sie durch kurzes Drücken der linken Maustaste, während sich der Mauszeiger innerhalb des verkleinerten Bildes befindet. Durch Drücken einer Taste (Maus oder Tastatur) wird er wieder verlassen.

Warnbox: Rechteckiges Feld, welches den Benutzer vor der Ausführung einer Funktion, welche Konsequenzen haben könnte warnt, oder ihn über geschehene Tatsachen in Kenntnis setzt. Es sind jeweils ein oder mehrere Antwortknöpfe vorhanden, welche mit der Maus betätigt werden können. Die Warnboxen werden in Kapitel 5.3 genauer erläutert und sind im Anhang D zusammengefasst.

Anhang D: Warnboxen

Dieser Anhang beinhaltet Fehlermeldungen und Warnungen, welche AUGUR in Form von Warnboxen anzeigen kann. Sie sind alphabetisch geordnet. Warn-Boxen haben ein oder mehrere Antwort-Knöpfe, welche mit der Maus bedient werden können. Text in solchen Antwort-Knöpfen ist im folgenden in Anführungszeichen gesetzt. Abb. D.1 zeigt als Beispiel eine Warnbox.

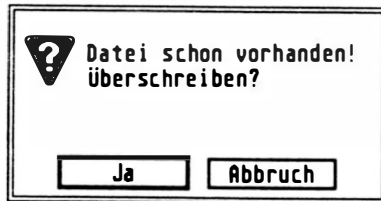


Abb. D.1. Beispiel einer Warnbox

ACHTUNG: Formatieren löscht alle Daten auf Ihrer Diskette!

Diese Meldung setzt Sie in Kenntnis, dass beim Formatieren alle Daten, die sich even-

tuell auf der Diskette befinden gelöscht werden.

Es müssen jedoch alle neuen Disketten vor dem ersten Gebrauch einmal formatiert werden.

Alle Rahmen sind gelöscht.

Es wurde versucht, einen Rahmen zu löschen, obwohl bereits alle Rahmen gelöscht sind.

Bibliothek voll

Es haben keine weiteren Prototypen mehr in der Bibliothek Platz. Vor dem nächsten Versuch müssen Sie eine frühere Version der Bibliothek laden oder die Bibliothek löschen. Mit "Abbruch" wird die Erkennung abgebrochen mit "Weiter" wird beim nächsten Rahmen fortgefahren. Dies hat jedoch nicht viel Sinn, da AUGUR nur bis zum nächsten unbekanntem Zeichen lesen kann.

Bibliothek zuerst sichern?

Sie haben Versucht, AUGUR zu verlassen,
ohne die erweiterte Bibliothek zu speichern.

Mit "Sichern" wird die Bibliothek gespeichert (es erscheint eine Dateiauswahl-Box), dann wird das Programm verlassen. Mit "Ende" wird das Programm verlassen, ohne die Bibliothek zu speichern und mit "Zurück" wird in das Programm zurückgekehrt.

Bitte Accessories zuerst schliessen!

AUGUR arbeitet nur, wenn alle Accessories geschlossen sind. Es könnte sonst zu Konflikten im Bildschirmaufbau kommen.

Das ROM-Modul steckt nicht richtig. Besser einstecken.

Das ROM-Modul konnte nicht richtig gelesen werden. Versuchen Sie die Daten (Text und Bibliothek, eventuell auch das Bild) im Computer noch abzuspeichern. Schalten Sie den Computer danach aus und überprüfen Sie, ob das ROM-Modul ganz eingesteckt ist. Oft hilft es auch, wenn Sie die vergoldeten Kontaktstifte mit einem weichen, fuselfreien Lappen und etwas Alkohol reinigen.

Datei-Lesefehler!

Beim Lesen einer Datei ist ein Fehler aufgetreten. Versuchen Sie es auf jeden Fall noch einmal, wenn möglich auch auf einem anderen Computer. Hilft dies auch nichts, so ist vermutlich die Diskette oder das Laufwerk defekt.

Datei schon vorhanden! Überschreiben?

Diese Meldung macht Sie darauf aufmerksam, dass Sie einen Dateinamen ausgewählt haben, der bereits existiert. Drücken Sie auf "Ja", so wird die alte Datei durch die neue ersetzt. Die alte Datei geht dabei also verloren. Drücken Sie "Abbruch", so wird der Vorgang abgebrochen, ohne dass die Datei gespeichert wird.

Diese Warnbox kann häufig auftreten, zum Beispiel wenn Sie eine erweiterte Bibliothek auf Diskette zurückschreiben.

Datei unwiderrufbar löschen?

Diese Meldung setzt sie in Kenntnis, dass beim Löschen einer Datei die entsprechenden Daten für immer verloren sind. Drücken Sie als Antwort "Ja", wird die ausgewählte Datei

gelöscht, mit "Abbruch" wird der Vorgang abgebrochen.

Datenübertragung unterbrechen?

Während dem Übertragen des erkannten Textes über die serielle Schnittstelle (Befehl "Text senden") wurde eine Taste gedrückt. Mit "JA" wird die Übertragung abgebrochen, mit "NEIN" wird sie weitergeführt.

Der Text wird gelöscht!

Vor dem Laden einer Bibliothek wird der Text im Arbeitsspeicher automatisch gelöscht. Mit dieser Meldung werden Sie davon in Kenntnis gesetzt. Wenn Sie damit einverstanden sind, drücken Sie "OK", ansonsten "Abbruch". Im zweiten Fall wird die Ausführung der Funktion abgebrochen.

Die Bibliothek ist leer!

Es wurde versucht, die Bibliothek abzuspeichern, oder den Lernmodus aufzurufen, ohne dass sich eine Bibliothek im Arbeitsspeicher befindet.

Der Lemmodus kann erst betreten werden, wenn eine Bibliothek geladen oder ein Text erkannt wurde. Dasselbe gilt für das Speichern der Bibliothek.

Dieses Programm läuft nur im Monochrom-Modus.

Diese Meldung erscheint, wenn AUGUR gestartet wird mit einem anderen als dem Schwarz-Weiss-Monitor (Monochrom) oder einem professionellen, hochauflösenden Monitor. Schliessen Sie den Monochrom-Monitor an, bevor Sie AUGUR wieder starten.

Disk-Fehler, kann Datei nicht öffnen

Lesefehler beim Lesen oder Schreiben auf Diskette. Geschieht dieser Fehler beim Abspeichern einer Datei, so verwenden Sie eine neue Diskette. Geschieht er beim Lesen, so nehmen Sie die Diskette aus dem Laufwerk, klopfen leicht daran und versuchen es nochmals.

Tritt dieser Fehler öfters auf, auch bei der Verwendung von Qualitätsdisketten (Double Sided, Double Density = 2S-DD), so könnte es sich um einen Fehler in Ihrem Diskettenlauf-

werk handeln. Setzen Sie sich in diesem Fall mit Ihrem Atari Händler in Verbindung.

Die Diskette in Laufwerk A lässt sich nicht formatieren!

Es ist entweder keine oder eine fehlerhafte Diskette in Laufwerk A (eingebautes Laufwerk).

Disk voll? Kann Datei nicht schreiben.

Die Diskette ist sehr wahrscheinlich voll. Nehmen Sie eine leere Diskette, um die gewünschte Datei abzuspeichern.

Dokument zu schief

Sie müssen das Dokument vor dem nächsten Versuch noch einmal scannen und dabei darauf achten, dass es weniger verdreht auf dem Scanner liegt. "Abbruch" beendet den Erkennungsvorgang, "Weiter" versucht den nächsten Rahmen zu lesen.

Es sind keine Rahmen mehr frei.

Es wurde versucht, mehr als 50 Rahmen einzuzichnen. Ist eine Textstruktur so kompliziert, dass mehr als 50 Rahmen nötig sind, so müssen Sie den entsprechenden Text in mehreren Schritten erkennen.

Falsches Bildformat!

Es wurde versucht ein Bild mit unzulässigem Format zu laden. AUGUR akzeptiert nur Bilder der Breite 1664 Bildpunkte (Hochformat) oder 2176 Bildpunkte (Querformat). Das Programm Scansoft kann durchaus andere Formate erzeugen. Um Bilder, welche mit Scansoft erstellt wurden in AUGUR verwenden zu können muss dort das OCR-Format ausgewählt werden.

Ich kann die Datei nicht finden

AUGUR kann die gewünschte Datei auf der Diskette oder Harddisk nicht finden. Überprüfen Sie, ob der richtige Dateiname angegeben wurde.

Interner Speicher voll

Der interne Arbeitsspeicher von AUGUR ist überlaufen. Der entsprechende Block kann nicht erkannt werden. Mit "Abbruch" wird der Erkennungsvorgang abgebrochen, mit "Weiter" wird beim nächsten Rahmen fortgefahren, dies ist jedoch sinnlos.

Kein Bibliotheks-Format!

Es wurde versucht mit der Funktion "Bibliothek laden" eine Datei zu laden, welche keine Bibliothek ist.

Kein Bild vorhanden!

Es wurde die Funktion "Bild sichern" ausgeführt, ohne dass sich ein Bild im Arbeitsspeicher befand (das Fenster in der Bildschirmmitte ist weiss).

Kein Text vorhanden!

Es wurde versucht, den Text anzuzeigen oder

abzuspeichern, ohne dass sich ein Text im Arbeitsspeicher befindet. Dies lässt sich am Eintrag unter dem Titel "Info" erkennen, dort steht, die Anzahl Zeichen sei 0. Text gelangt nur während des Erkennungsvorganges in den Arbeitsspeicher.

Keine Zeilen gefunden

Es handelt sich um eine schlechte Vorlage, Grafik oder es ist gar kein Bild vorhanden. "Abbruch" unterbricht den Erkennungsvorgang, "Weiter" fährt mit der Erkennung des nächsten Rahmens fort.

Es kann aber auch sein, dass sich am linken oder Rechten Bildrand ein dünner, kaum sichtbarer vertikaler schwarzer Strich befindet (es könnte zum Beispiel ein Schatten der Vorlage sein), welcher das Erkennen verhindert. Legen Sie einen Rahmen um den Text und versuchen Sie es noch einmal. Oft bringt auch das Ausführen der Funktion "Linien entfernen" Abhilfe.

Nicht genügend Speicher vorhanden.

Es ist zuwenig Arbeitsspeicher für AUGUR vorhanden. Abhilfe bringt das Entfernen der

Accessories und residenten Programme. Beachten Sie aber, dass AUGUR nur auf Computern mit mindestens 1 MByte Arbeitsspeicher läuft. Auf dem Atari 520ST kann AUGUR also nicht betrieben werden.

Diese Warnbox kann ebenfalls erscheinen, wenn versucht wird eine Bibliothek zu laden, welche nicht mehr im Arbeitsspeicher platz hat. Dies kann vorkommen, wenn besagte Bibliothek auf einem Computern mit mehr Speicherplatz aufgebaut worden ist. Je grösser der Speicherplatz des Computers, desto grösser ist auch die maximale Grösse welche eine Bibliothek annehmen kann.

Scanner antwortet nicht. Kabel und Schalter überprüfen.

Die Kommunikation mit dem Scanner funktioniert nicht. Mögliche Ursachen sind: Der Scanner ist nicht mit dem Atari verbunden, einer der Stecker hat sich gelöst, der Scanner ist nicht eingeschaltet oder der falsche Scanner ist angewählt.

Seriewerk abbauen?

Sie haben während der Seriewerk eine Taste

gedrückt. Mit "Ja" wird die Seriearbeit abgebrochen, mit "Nein" wird weitergefahren.

Sie haben keinen Dateinamen angegeben.

In der Objektauswahl-Box wurde kein Dateiname angegeben. Sie müssen vor dem Speichern oder Laden einer Datei den gewünschten Namen in die Auswahlzeile schreiben oder mit der Maus einen Namen im Inhaltsverzeichnis auswählen. Die Dateiauswahl Box wird in Kapitel 5. "Vor dem Anfang" genau beschrieben.

Text anhängen an:

Diese Box erscheint nach dem Starten der Seriearbeit. Auf der zweiten Zeile wird der Name der Datei angezeigt, an welche der Text jeweils automatisch angehängt werden wird. Sind Sie damit einverstanden, so klicken Sie "OK", andernfalls klicken Sie "Abbruch" und die Seriearbeit wird nicht gestartet.

~Vollen Sie im zweiten Fall den Dateinamen ändern, so müssen Sie die erste Seite mit "Text erkennen" lesen und anschliessend mit "Text sichern" in einer neuen Datei speichern oder mit "Text anhängen" an eine bestehende Datei anhängen. Beim erneuten Starten der Seriearbeit wird diese Datei genommen werden, um

weiteren Text anzuhängen.

Text löschen?

Die Schrifterkennung ist gestartet worden, wobei sich bereits ein Text im Arbeitsspeicher befindet. Mit "Ja" wird dieser Text gelöscht, mit "Nein" wird der neu erkannte Text diesem hinzugefügt und mit "Abbruch" wird die Erkennung abgebrochen.

Text zuerst sichern?

Sie haben Versucht, AUGUR zu verlassen, ohne den erkannten Text zu speichern. Mit "Sichern" wird der Text gespeichert (es erscheint eine Dateiauswahl-Box), dann wird das Programm verlassen. Mit "Ende" wird das Programm verlassen, ohne den Text zu speichern und mit "Zurück" wird in das Programm zurückgekehrt.

Unkomprimierbares Bild. Datenverlust!

Bilder werden vor dem Speichern immer komprimiert. In diesem Fall wurde versucht, ein

Bild, welches sich nicht komprimieren lässt, zu speichern. Da wegen der grossen Datenmenge der Bildarbeitsspeicher selbst zur Kompression benützt werden muss, wurde das Bild dabei zerstört.

Dieser Fehler ist in der Praxis noch nie aufgetreten.

Unlesbare Schrift

Das Dokument hat zu schlechte Qualität und kann von AUGUR nicht gelesen werden.

Vor "Seriearbeit" Datei mit "Text sichern" eröffnen.

Bevor die Seriearbeit gestartet werden kann, muss die erste Seite der zu erkennenden Dokumente mit "Text sichern" gespeichert werden. Damit wird AUGUR mitgeteilt, an welche Datei der erkannte Text während der Seriearbeit angehängt werden muss.

Zu viele Zeichen

Der Arbeitsspeicher für die erkannten Texte ist voll. Löschen Sie den Text und speichern Sie

ihn gegebenenfalls vorher ab. Sie haben als Antwort "Abbruch" und "Weiter" zur Verfügung, wobei nur "Abbruch" sinnvoll ist.

Zu viele Zeilen

Der Zeilenpuffer von AUGUR ist überlaufen. Brechen Sie den Erkennungsvorgang ab ("Abbruch") und unterteilen Sie den entsprechenden Rahmen in zwei kleinere Rahmen mit weniger Zeilen.

Anhang E: ASCII-Tabelle

In diesem Anhang ist der ASCII-Code (American Standard Code for Information Interchange), sowie der erweiterte Atari-ASCII-Code (Sonderzeichen) tabelliert.

In diesem Code werden Texte von AUGUR abgespeichert. Ist das 1st Word Plus Format aktiv, so kommen noch Formatinformationen hinzu.

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
2	!	"	#	\$	%	&	'	()	*	+	,	-	.	/	
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	Δ
8	Ç	ü	é	â	ä	à	ä	ç	ê	ë	è	î	ï	Ë	Ä	Å
9	È	æ	Ë	ô	ö	ò	ü	ü	ÿ	ö	ü	ç	£	¥	ß	f
A	á	í	ó	ú	ñ	ñ	á	ó	ú	ÿ	ö	ü	ç	£	¥	ß
B	ä	ö	ð	ø	æ	æ	á	ó	ú	ÿ	ö	ü	ç	£	¥	ß
C	ij	Y	X	1	7	8	9	0	1	2	3	4	5	6	7	8
D	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5
E	α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο	π
F	≡	±	≤	≥	∫	∫	∫	∫	∫	∫	∫	∫	∫	∫	∫	∫

Steuercodes:
 0A: Line Feed
 0C: Form Feed
 0D: Carriage Return

Abb. E.1. Der Atari-Zeichensatz

Anhang F: Mögliche Fehlerquellen

Wichtig: Wenn Sie eine Vorlage haben, von der Sie glauben, dass sie für AUGUR lesbar sein sollte, die aber auch nach Abklärung der untenstehenden Punkte, nicht gelesen wird, so speichern Sie das entsprechende Bild bitte auf eine Diskette und schicken Sie diese mit einem möglichst ausführlichen Kommentar an:

Marvin AG
Abt. Support AUGUR
Friesstrasse 23
CH-8050 Zürich.

Wir werden bemüht sein, solche Fehler in der nächsten Version von AUGUR zu beheben.

Der Bildschirm bleibt dunkel. Prüfen Sie ob der Bildschirm am Computer angeschlossen und auch eingeschaltet ist. Vergewissern Sie sich, dass das AUGUR-Steckmodul nicht verkehrt eingesteckt ist

(richtig: Beschriftung nach oben) und richtig festsetzt.

AUGUR stürzt ab. Schalten Sie den Computer für einige Sekunden aus und versuchen Sie es noch einmal. Es kann sein, dass im System gewisse Sachen "versteckt" waren, die ein einwandfreies Funktionieren von AUGUR verhindert haben.

Reinigen Sie die vergoldete Kontaktleiste des AUGUR-Steckmoduls mit Alkohol und einem fuselfreien Lappen.

Es kann aber auch sein, dass gewisse residente Programme oder Accessories unsauber programmiert sind und sich deshalb nicht mit AUGUR "vertragen". Entfernen Sie alle Accessories und residenten Programme und versuchen Sie es noch einmal. Booten Sie dazu den Computer mit der mitgelieferten Diskette.

Des weiteren besteht die Möglichkeit, dass sich AUGUR nicht mit externer oder interner Hardware verträgt, welche sich nicht nach den Normen von Atari richtet. Bauen Sie alle zusätzliche Hardware aus Ihrem Computer aus und entfernen Sie

alle Peripheriegeräte ausser dem Scanner und versuchen Sie es noch einmal.

Der Hawk CP14 Scanner lässt sich nicht starten: Überprüfen Sie ob sich noch Papier im Scanner befindet (rote Lampe "Paper Empty" auf dem Bedienungsfeld darf nicht leuchten).

Schriftzeilen werden ausgelassen oder nicht getrennt: Dieses Problem kann bei eng beieinanderliegenden Schriftzeilen auftreten. AUGUR vermag dann die Zeilen nicht zu trennen und hält zusammengebliebene Schriftzeilen zum Teil für Grafik und lässt diese weg. Oft schafft das hellere Scannen des dokumentes Abhilfe.

Alleinstehende Schriftzeichen werden ausgelassen. Dies kann unter Umständen Absicht von AUGUR sein. Alleinstehende, also nicht in einer Textstruktur integrierte Zeichen, welche einen zu kleinen relativen Schwarzflächenanteil besitzen werden ausgelassen. Dies hat sich als sehr

nützlich erwiesen, um zu vermeiden, dass Heftklammern und ähnliches versucht wird zu erkennen.

Abhilfe kann bei Bedarf manchmal geschaffen werden, in dem die isolierten Zeichen separat eingerahmt werden. Diese werden von AUGUR dann nicht mehr als "Ausreisser" betrachtet.

Eine Datei erscheint nicht in der Objektauswahl-Box. Dies könnte folgende Ursachen haben:

- die falsche Diskette ist eingelegt
- in der Indexzeile steht der falsche Pfad (Laufwerk) oder Extender
- die gesuchte Datei befindet sich in einem Ordner.

Im Notfall löschen Sie die Indexzeile und klicken Sie in den Titelbalken des Inhaltsverzeichnisses. Es werden daraufhin alle Dateien angezeigt, welche sich zuoberst (nicht in einem Ordner) im Laufwerk A befinden.

Beim Laden eines Bildes kommt nur "Rauschen". Es ist möglich, dass Sie

kein Bild, sondern eine andere Datei (Bibliothek, Text etc.) geladen haben. AUGUR glaubt in diesem Fall, es handle sich um ein unkomprimiertes Bild und zeigt die entsprechende Datei als Bild an, was natürlich keinen Sinn ergibt.

Anhang G: Beispiel-Dokumente

In diesem Anhang befinden sich drei Dokumente als Beispiel für die Schrifterkennung. Die Zugehörigen Bibliotheken, sowie bereits gescannte Bilder befinden sich auf der mitgelieferten Beispiel-Diskette.

1. Brief

Ein Beispiel in klassischer Schreibmaschinenschrift, fixed spaced.

2. Zürich

Ein Dokument in Frakturschrift (proportional). Dieses Beispiel demonstriert die hohe Lernfähigkeit von AUGUR.

3. Testbericht

Ausschnitt aus einem Testbericht über AUGUR. Hier wird die Fähigkeit, Rahmen zu zeichnen und Proportionalschrift zu erkennen

gebraucht.

Register

Das Register kann erst erstellt werden, wenn alle Korrekturen im Handbuch abgeschlossen sind.

(der Autor)