



Handschriften erkennen

Leseprobe

Gute Erkennungsleistungen bieten Programme wie FineReader oder Recognita schon lange – bei Handschriften stieg OCR-Software bisher aus. SoftWriting 3.2 liest sogar **handschriftliche Texte**.

DR. RALF THOMAS GÖLLNER

Mit SoftWriting 3.2 Lite von CharacTell liegt unserer Heft-CD eine besondere Texterkennungs-Software bei: Sie liest und erkennt handschriftlich geschriebene Texte. Mit diesen Tipps und Tricks zeigen wir, wie Sie das Beste aus der ICR (Intelligent Character Recognition) herausholen. Aber Vorsicht: Hieroglyphen kann auch eine intelligente Software (noch) nicht entziffern!

1 Gut vorbereitet

Während OCR-Programme (Optical Character Recognition) wie FineReader weniger gute Vorlagen verzeihen, ist SoftWriting auf einwandfreies Ausgangsmaterial angewiesen.

Verwenden Sie weißes, unliniertes Papier und beschreiben Sie es nur einseitig. Bei doppelseitiger Beschriftung sollte das Papier mindestens 80 g/m² wiegen. Sonst scheint die Rückseite durch und verfälscht das Ergebnis.

Schreiben Sie am besten mit einem schwarzen Stift, der gleichmäßige Linien zieht, und drücken Sie gleichmäßig auf – Unterbrechungen oder deutliche Aufhellungen einer Buchstabenlinie quitiert SoftWriting mit Fehlern.

Die Spitze Ihres Schreibstiftes sollte dünn sein, so dass SoftWriting auch bei schneller Schreibweise sich ähnelnde Buchstaben unterscheiden kann: Sonst wird aus einem *e* schnell ein *c*.

2 Stilfragen

Schreibschriften mit verbundenen Buchstaben erkennt SoftWriting schlecht. Deshalb müs-

sen Sie Ihren Schreibstil ändern und die Buchstaben getrennt zu Papier bringen.

Sie können Groß- und Kleinbuchstaben verwenden, bessere Erkennungsraten erzielen Sie aber mit großen Druckbuchstaben. Für die Weiterverarbeitung ist das irrelevant, die Software wandelt sie anschließend meist sinnvoll um. Achten Sie darauf, dass sich die Buchstaben eines Wortes nicht berühren. In der Regel kann SoftWriting diese nicht mehr erkennen.

Die Abstände sollten auch nicht zu groß sein, sonst deutet die ICR sie nicht als zusammenhängendes Wort. Schreiben Sie groß genug, um die Besonderheiten Ihrer Schreibschrift zur Geltung zu bringen. Lassen Sie zwischen den Zeilen ausreichend Platz, sie dürfen sich keinesfalls berühren. Achten Sie auf deutliche Umlaute, diese kann sonst Probleme bereiten. SoftWriting ist gelehrt und wird mit häufiger Anwendung zuverlässiger.

3 Erfassen-Optionen wählen

Die richtigen Scan-Einstellungen erhöhen die Leseleistung.

Um mit der ICR arbeiten zu können, richten Sie zunächst einen neuen Benutzer (*Schreiber*) ein. Die Lite-Version unterstützt nur ein Anwenderprofil. Folgen Sie dem *LearnWizard* bis zur Auswahl der Bildquelle. SoftWriting unterstützt den Scanner- wie den Batch-Modus. Das Scannen im Netzwerk funktionierte bei unseren Tests leider nicht, so dass ein am PC angeschlossener Scanner notwendig war.

Öffnen Sie die Scanner-Optionen. Die ideale Auflösung hängt vom verwendeten Schreibstift ab. Meist erzielen Sie mit 300 dpi Auflösung bessere Ergebnisse.



AUF CD

CharacTell SoftWriting 3.2 Lite finden Sie unter der Rubrik *Software/Vollversionen*.

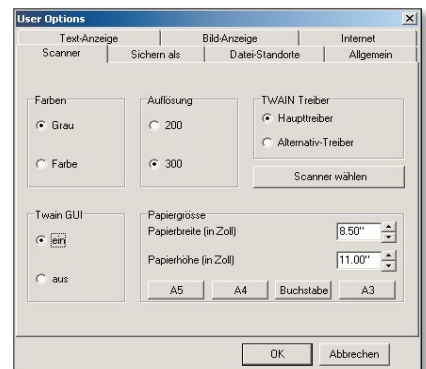


Verwenden Sie den vom Scanner-Hersteller mitgelieferten TWAIN-Treiber, indem Sie *Haupttreiber* anklicken.

Setzen Sie *TWAIN GUI* (Graphical User Interface) auf *ein*. So können Sie auf die Einstellungen des TWAIN-Treiber vor dem Scanvorgang zugreifen und Helligkeit und Kontrast anpassen.

Die Papiergröße überprüfen Sie auch im TWAIN-Treiber, denn der *Alternativ-Treiber* von SoftWriting schneidet eine DIN A4-Seite etwa beim Letter-Format ab. Die letzten Zeilen der Seite fehlen dann.

Legen Sie die Seite richtig ausgerichtet ein, so dass sie nicht auf dem Kopf steht, und erfassen Sie sie.



HAUPTTREIBER ERSTE WAHL: die Einstellungen in den Scanner-Optionen

Fügen Sie anschließend so viele Seiten hinzu, bis Sie etwa 400 bis 500 Wörter gescannt haben. Zuviel Text lehnt SoftWriting ab und fordert Sie so lange zum Löschen auf, bis eine für die Texterkennung brauchbare Wortanzahl vorliegt.



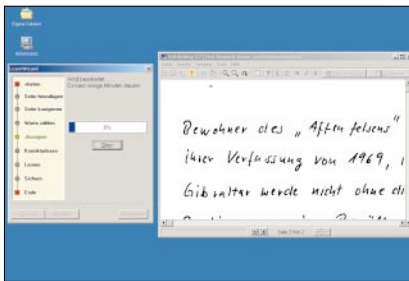
DIE TEXTMENGE für den Lernvorgang reicht hier aus.



4 Lesen bildet

Wenn der *LearnWizard* den Text erkennt, dauert dies abhängig von Ihrer Schrift, dem Textumfang und der Leistung Ihres Computers eine Weile.

Wenn sich beim Erkennen auf Ihrem Bildschirm mehrere Minuten nichts regt, ist das Programm nicht abgestürzt. Der *LearnWizard* liegt dann verborgen hinter Ihrem *SoftWriting*-Fenster. Sie bringen ihn zum Vorschein, indem Sie auf das *Verkleinern/Maximieren*-Symbol in der Fensterleiste klicken und das verkleinerte Fenster verschieben. Jetzt können Sie auch den Lern- und Erkennungsfortschritt verfolgen.



AUFGEDECKT: Per *Verkleinern/Maximieren*-Symbol bringen Sie den *LearnWizard* zum Vorschein.

Ist die erste Phase der Erkennung und des Lernens abgeschlossen, *beurteilt* *SoftWriting* Ihre Handschrift. Es meldet, ob die Erkennungsrate durchschnittlich ist oder unter dem Durchschnitt liegt. In letzterem Fall bietet es die Möglichkeit, weiterzumachen oder den Lernprozess abzubrechen. Wenn Sie deutlicher schreiben können, versuchen Sie einen neuen Durchgang. Sonst bleibt Ihnen nichts anderes übrig, als weiterzumachen und den Text zu korrigieren.

5 Handarbeit

Das Wörterbuch kennt anfangs nicht jedes Wort, auch Zahlen und Personennamen sind ihm fremd. Hier hilft *Training*.

Findet die Software nach dem ersten Lernprozess ein Wort nicht im Wörterbuch, wird es rot dargestellt. Bei blau geschriebenen Wörtern ist sie sich nicht sicher, ob die Deutung der Algorithmen mit dem Eintrag des Lexikons übereinstimmt.

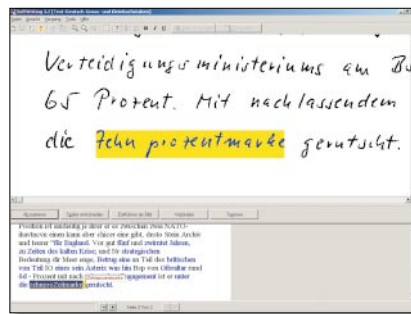
Durch den Korrekturvorgang bringen Sie aber nicht nur der *ICR-Engine* Ihre Handschrift bei, sondern ergänzen auch

das Wörterbuch durch neue Wörter. Dadurch wird die Erkennungsleistung doppelt gesteigert und der Wortschatz von *SoftWriting* immer umfangreicher. Mit der Zeit erlernt es auch Zahlen. Vorsicht: Tippfehler an dieser Stelle wirken sich langfristig fatal aus – Sie haben ein fehlerhaftes Wörterbuch, denn die Einträge werden ungeprüft übernommen.

TIPP Die Korrekturen können Sie vornehmen, indem Sie die falsch erkannten Wörter überschreiben. Am schnellsten springen Sie mit *[Strg]* und *[Pfeil]*-Links bzw. *[Pfeil]*-Rechts von Wort zu Wort, können aber auch die *Maus* verwenden.

Verlassen Sie sich nicht auf die farbige Darstellung, auch scheinbar richtig erkannte Wörter können „*sinnvoll*“ falsch sein, wenn sie beispielsweise getrennt erkannt werden. So bedeutet zum Beispiel *über gehen* etwas anderes als *übergehen*.

Im geteilten Fenster sehen Sie oben die graphische Darstellung Ihrer Vorlage, wobei das gerade aktive Wort gelb unterlegt ist. Dies hilft Ihnen, richtig zu vergleichen und zu manövrieren. Hier können Sie prüfen, ob bei der Erkennung ein zusammengehörendes Wort getrennt oder nur falsch erkannt wurde. Im ersten Fall gehen Sie bis zum ersten Wortteil und klicken dann auf *Verbinden*. *SoftWriting* erweitert die gelbe Markierung nach rechts und aktualisiert sofort das Ergebnis im unteren Fensterabschnitt.

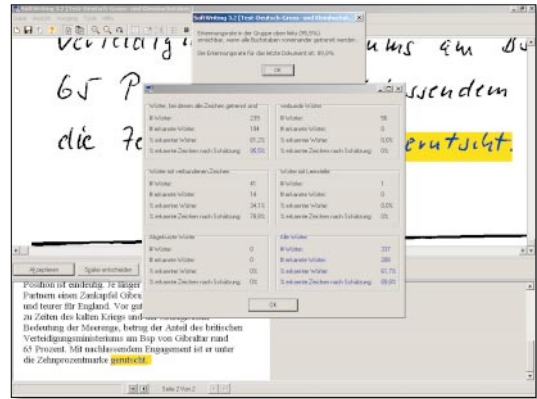


KLEBBAND: Verbinden Sie getrennt erkannte Wortteile zu einem zusammenhängenden Wort.

In vielen Fällen bietet *SoftWriting* mit einem sich automatisch öffnenden Drop-down-Feld eine Auswahl aus dem Wörterbuch. Ist das richtige Wort dabei, können Sie mit der *[Pfeil-Unten]*-Taste zum richtigen Wort springen. Korrekt erkannte Wörter bestätigen Sie mit der *[Enter]*-Taste.

6 Lebenslanges Lernen

Haben Sie den Text vollständig durchgearbeitet und alles korrigiert, können Sie eine detaillierte Statistik abrufen. Diese dient Ihnen als Kontrolle, wie Sie Ihren Schreibstil anpassen können.



LASST ZAHLEN SPRECHEN: Über die Erkennungsstatistik können Sie die Leseleistung erhöhen.

Die Aufschlüsselung nach *Wörter*, bei denen alle Zeichen getrennt sind, *Wörter* mit verbundenen Zeichen oder *verbundene Wörter* zeigt, mit welchen typischen Schreibgewohnheiten das Programm Schwierigkeiten hat. Bei einer erneuten Erkennung sollten Sie auf diese Besonderheiten achten, um die Erkennungsrate zu steigern. P K

UPGRADE

Im Unterschied zu *SoftWriting 3.2 Lite* unterstützt die große Variante frei definierbare Abkürzungen für den handschriftlichen Text. *SoftWriting* wandelt dann ein handgeschriebenes *z B* in *zum Beispiel* um. Zudem übernimmt es Zeichnungen als Grafik in den Text und bietet die sinnvolle Option, gescannte Bereiche vom Erkennungsvorgang auszuschließen bzw. zu löschen. Dies ist hilfreich, weil auch durchgestrichene Zeichen oder unerwünschte Skizzen ausgeblendet werden können. Dadurch arbeitet das Programm auch schneller. Die große Variante unterstützt neben Deutsch auch Englisch, Spanisch, Portugiesisch, Französisch, Italienisch und Niederländisch. Für einen Preis von 19,95 US-Dollar erhält der Benutzer ein Upgrade mit den erweiterten Funktionen. Weitere Informationen erhalten Sie, indem Sie auf *Hilfe/Aktualisierungs-Information* klicken oder unter www.charactell.com.