

Wissenswertes und Praxistipps zur
Vektorisierung

Version:	1.0
Angefertigt von:	Wilfried Klemmer R.+S. Consult GmbH Schänzchen 5 53949 Dahlem Tel.: 02447 911048
Status:	Freigabe
Dateiname:	Vektorisierung.doc

Wissenswertes und Praxistipps zur Vektorisierung

(Dipl.-Ing. Wilfried Klemmer Juni/2005)

Begriffliche Grundlagen

Unter einer „Vektorisierung“ versteht man im allgemeinen GIS-Sprachgebrauch die Herstellung von Vektordaten (meistens auf der Basis von Rasterdaten). Bevor weitere Ausführungen hierzu gemacht werden, muss der Unterschied zwischen Raster- und Vektordaten für die GIS-Anwendung deutlich gemacht werden.

Grundsätzlich handelt es sich bei beiden Datenformaten um Strukturen zur Darstellung von Grafik. In diesem thematischen Zusammenhang soll davon abgesehen werden, dass Raster- und Vektordaten natürlich auch zur Darstellung alphanumerischer Daten verwendet werden können. Rasterdaten sind nichts anderes als ein Mosaik, das aus regelmäßig angeordneten, gleich großen Mosaiksteinen besteht. Ein Beispiel für diese Technik ist das Fernsehbild. Tritt man nahe genug an die Bildfläche heran, sieht man ihre Mosaikstruktur, die aus einer gewissen Entfernung betrachtet, den Eindruck eines Bildes vermittelt.

Das absolut gleiche Prinzip wird bei Rasterdaten verwendet. Einen „Mosaikstein“ nennt man dort ein Pixel. Je dichter die Menge an Pixeln eines Bildes ist, desto konturenschärfer erscheint es und desto größer ist die Datenmenge, die dieses Bild trägt. Als Maß für die „Konturenschärfe“ nimmt man die Anzahl der Pixel, die sich entlang einer Strecke von einem Inch (= 2,54 cm) befinden. Dieses Maß wird als „Auflösung“ bezeichnet und mit dpi (=Dots per Inch = Anzahl Pixel pro 2,54 cm) gemessen.

Rasterdaten werden als ganze Datei adressiert. Mit dem Aufruf einer Datei wird das entsprechende Mosaikbild auf dem Bildschirm reproduziert. Rastereditoren können zusätzlich auch noch die einzelnen Pixel in ihrer Farbgebung beeinflussen, so dass bildtechnisch Funktionen wie Einfärben von Bereichen, Kopieren, Löschen von Bereichen des Bildes, Zeichnen von Linien und Texten u.s.w. erreicht werden.

Vektordaten sind ganz anders aufgebaut. Sie bestehen in ihren Grundkonstrukten eigentlich nur aus Punkten und Verbindungen zwischen den Punkten (Linien) und entsprechen damit ganz der mathematischen Definition eines Vektors - einer durch zwei Punkte bestimmten, gerichteten Strecke (im Raum oder auch nur in der Ebene). Ausgehend von diesen Grundkomponenten baut man noch Polygonlinien (mehrere aneinander gekettete Linien), Flächen (geschlossener, sich nicht schneidender Polygonzug) und Texte auf.

Bildlich kann man sich Vektordaten als Konturzeichnungen vorstellen. Das Besondere an den Vektordaten liegt darin, dass sie über die Dateiansprache hinaus noch differenzierter adressiert werden können. Mit Hilfe der oben definierten Grundkonstrukte (Punkt, Linie, Fläche, Text) können Objekte definiert werden, die dem Abbild eines Objektes der realen Welt (z.B. im Grundriss) entsprechen. Damit hat man genau die Möglichkeiten, die man zum Zeichnen von Karten, bzw. die man zum Aufbau eines grafischen Systems benötigt. Anstelle der Objekte der realen Welt platziert man in maßstäblicher Weise deren grafische Symbole, die punkt-, linien-, flächen oder textförmig sind. Jedes dieser Objekte ist für das grafische System einzeln ansprechbar, d.h., man kann ihm bestimmte Bezeichnungen und Attribute zuweisen und über diese Daten sind diese Objekte dann wiederum umfassend auswertbar.

Erst Vektordaten (und ihre direkte Adressierbarkeit in einem grafischen DV-System) machen eigentlich ein Geografisches Informationssystem (GIS) erst zu dem, was es seiner Definition nach sein muss. Vektordaten bilden damit die absolute (notwendige) Grundlage für ein GIS. Kein Wunder also, dass der Erzeugung von Vektordaten eine hohe Bedeutung zukommt.

Worin besteht die Schwierigkeit?

Vom Prinzip her ist die Erstellung von Vektordaten überhaupt nicht schwierig. Alle grafischen Systeme haben einen recht großen Vorrat an Funktionalität, mit denen Vektordaten erzeugt werden können. Das Problem besteht „nur“ im Aufwand, der erforderlich ist, um diese Daten zu erzeugen. Da Vektordaten erfasst und zu maßstäblichen Objekten in einem grafischen System werden sollen, müssen sie nach bestimmten Verfahren erzeugt werden. Diesen Vorgang der Erzeugung bezeichnet man als grundsätzlich als „Digitalisieren“. Synonym wird auch „Vektorisieren“ als Begriff genannt.

Für die digitale Erfassung von Vektordaten bestehen im Prinzip zwei Möglichkeiten:

- Konstruktion oder
- Bildschirmdigitalisierung.

Das Konstruieren ist der mühsamste Weg zur Erzeugung der Vektordaten. Mit Hilfe der Konstruktionsfunktionen eines grafischen Systems wird dabei das entsprechende Objekt maßgerecht aufgebaut. Dieses Verfahren beinhaltet die höchste Genauigkeit, aber auch den größten manuellen Aufwand. Vorteilhaft ist: Maße, die später aus dieser Konstruktion abgegriffen werden, sind exakt.

Die Bildschirmdigitalisierung lässt sich etwas leichter durchführen. Üblicherweise geht man (heute) davon aus, dass man als Basis des Digitalisierens eine georeferenzierte Rasterdatei oder eine Vektordatei im System geladen hat, auf deren Grundlage man die Punkte, die zu digitalisieren sind, mit der Zieleinrichtung des grafischen Systems (Fadenkreuz, Pfeil o.ä.) zur Deckung bringt und dann digital erfasst (Bildschirmdigitalisierung). Die früher übliche Digitalisierung mit einem Digitalisiertisch ist heute kaum noch anzutreffen. Anhand der abdigitalisierten Punkte wird während oder nach dem Erfassungsvorgang die Objektbildung vollzogen. Dieses Verfahren hat Ähnlichkeiten mit dem „Hochzeichnen“ eines Bildes auf Transparentpapier. Da es direkt auf einer Vorlage durchgeführt wird, ist es anschaulicher und schneller durchzuführen als die Konstruktion. Allerdings hängt die Genauigkeit von vielen Einflussfaktoren ab.

(Das vermessungstechnische Erfassen von Punkten und deren Überleitung in ein GIS entspricht als Grundlage genau den Vektordaten und wird deshalb nicht gesondert erwähnt.)

Anhand eines praktischen Beispiels soll dies deutlich gemacht werden. Erstellt man einen Bebauungsplan durch Konstruktion, kann man die durch die Konstruktion eingebrachte Geometrie exakt auswerten. Eine definierte Abstandslinie zu einer Grenze hat dann auch genau den Sollabstand. Wäre diese Linie auf der Basis eines Bebauungsplanes per Bildschirmdigitalisierung erfasst worden, so hätte der Abstand mit einiger Sicherheit Abweichungen zum Sollmaß. Diese Ungenauigkeit resultierte im wesentlichen aus dem Verzug des Basisplans, der Genauigkeit des Scannens des Basisplans, den nicht exakt getroffenen Zielpunkten bei der Digitalisierung, der Ungenauigkeit der Georeferenzierung des

Basisplanes und vieles mehr oder durch die Ungenauigkeit der örtlichen Aufnahme, wenn vermessene Punkte die Basis gewesen wären.

Wer also exakte Maße aus einem grafischen System herleiten muss, muss Verfahren anwenden, die letztlich auf konstruktionstechnischen Prinzipien beruhen. Wer nur eine bestimmte Genauigkeit benötigt, kommt mit der Bildschirmdigitalisierung aus.

Natürlich gibt es auch Bildschirmdigitalisierungen, die exakt definierte Punkte „einfangen“ können, weil die Grundlage aus Vektordaten stammt (Konstruktion oder Vermessungspunkte). Um hier die wirklich wichtigen Prinzipien heraus zu arbeiten, wird nur auf die beiden klassischen Methoden eingegangen.

Bei der Erzeugung von Rasterdaten hat man die oben erwähnten Probleme nicht. Ein Plan wird einfach über einen Scanner erfasst und als Datei abgespeichert. Dieser Vorgang ist sehr schnell durchführbar und kostengünstig. Hieraus leitet sich ein Gedankengang ab, der nahe liegend ist:

Wenn doch ein Mensch auf der Basis von Rasterdaten (durch die Bildschirmdigitalisierung) Vektordaten erfassen kann und die Erzeugung von Rasterdaten so kostengünstig ist: warum kann man dann nicht eine Software entwickeln, um den Vorgang der Bildschirmdigitalisierung automatisch durchzuführen?

Der enge Begriff der „Vektorisierung“

Häufig wird der Begriff der „Vektorisierung“ eng ausgelegt und meint genau die Konvertierung von (Teilen von) Rasterdaten in Vektordaten. Das Problem besteht darin, dass die Rasterdaten nur als Datei (im Ganzen) für eine Software erkennbar ist. Wie dargelegt, gibt es keine einzeln adressierbaren Objekte. Einer automatischen Vektorisierung muss es also gelingen, das Rasterbild fachgerecht zu interpretieren, aus der Ansammlung von Mosaiksteinen Objekte zu erkennen und diese Objekte dann in DV-technisch adressierbare Objekte im Vektorformat umzuformen.

Mit dem recht harmlos klingenden Begriff „Vektorisierung“ ist damit u.a. der äußerst komplexe Problembereich der Bild-, bzw. der Mustererkennung verbunden. Angesichts der Dichte der Darstellungen, der Qualität der vorhandenen Karten und des derzeitigen Standes der Forschung zur Mustererkennung hat man hier einen Problembereich vorliegen, der hinsichtlich seiner Lösung noch in den Kinderschuhen steckt.

Je besser ein Planwerk hinsichtlich seiner Darstellungsqualität ist, je geringer die Zahl und je einfacher die Struktur der zu erfassenden Objekte ist und je größer die Freiräume zwischen den Darstellungen in der Karte sind, desto eher gelingt eine Vektorisierung. Wenn hier von „Gelingen“ gesprochen wird, setzt man allerdings relativ bescheidene Prozentraten voraus.

Kurzum: eine automatische Vektorisierung mit hohen Erkennungsraten ist heute immer noch die Ausnahme. Selbst wenn gute Erkennungsraten erreicht werden, bleibt noch das Problem, dass man die Objekte identifizieren muss, die nicht, falsch oder unzureichend erkannt wurden und diese Objekte dann in manueller Form erfassen, bzw. korrigieren muss. Dieser Aufwand ist beträchtlich. Aus wirtschaftlichen Gründen kommt die automatische Vektorisierung in der Praxis noch recht selten zum Einsatz.

Praxistipps

Für die Praxis kann nur zu folgender Vorgehensweise geraten werden:

- Wenn für ein GIS eine Datenerfassung gemacht werden muss, prüfen Sie zunächst, welches Maß der Adressierung Sie für die Objekte brauchen.
Nicht immer benötigt man wirklich Objekte, die man dediziert ansprechen können muss. Oft genügt auch der einfache Plan als Anschauung im Hintergrund. In diesen Fällen ist das Scannen und Georeferenzieren die günstigste Alternative.
Typisches Beispiel: bestehende Bauleitpläne. Die ganzen Grenzen und Gebiete der Bauleitpläne zu erfassen, wäre ein sehr großer Aufwand. In den allermeisten Fällen genügt es, wenn die Bauleitplanung passgerecht im Hintergrund aufgerufen und zusammen mit den anderen Themen angeschaut werden kann.
- Benötigt man dedizierte Objekte, prüfen Sie, ob eine Erfassung über Bildschirmdigitalisierung (von der Genauigkeit her) ausreicht.
Da viele einen ähnlichen Darstellungsstand wie in den bestehenden Karten auch zukünftig im GIS wünschen, ist die Bildschirmdigitalisierung oft das probate Mittel für eine kostengünstige Vektorisierung. Bedenken Sie, dass eine Bildschirmdigitalisierung nicht zwangsläufig nur auf der Basis von Rasterdaten erfolgen muss. Auch Vektordaten können Basis für diese Digitalisierung sein. Der besondere Vorteil ist: diese Daten können in einfacher (Koordinaten) oder auch komplexer Form (Linien, Flächen) übernommen werden, so dass auch die Objektbildung für GIS weiter vereinfacht wird.
- Müssen exakte Maße bei einer GIS Auswertung abgeleitet werden können, helfen nur Konstruktionen bei der Objektbildung. Einzige Ausnahme: die Maße sind örtlich hinreichend genau aufgemessen und stehen als Punktkoordinaten im GIS zur Verfügung. Dann werden diese Punkte bei der Bildschirmdigitalisierung „eingefangen“ und für die Digitalisierung verwendet. In bestimmten Fällen ist auch die Kombination zwischen Bildschirmdigitalisierung und Konstruktion das wirtschaftlich und fachlich beste Mittel. Bedenken Sie aber auch, dass Bildschirmdigitalisierung eine Konstruktion nicht ersetzen kann! Wird hier vorab nicht genau festgestellt, welche Art der Vektorisierung erforderlich ist, ist u.U. eine ganze Bildschirmdigitalisierung wertlos. Typisches Beispiel: Digitalisierung von Gebäudeplänen zum Zwecke der Erweiterung der Konstruktionspläne. Der Architekt leitet aus dieser Unterlage Maße her. Was geschieht, wenn Fertigteile aufgrund eines Planes, der auf der Basis der Bildschirmdigitalisierung erstellt wurde, hergeleitet werden, kann sich jeder vorstellen.
- Lassen Sie sich niemals auf das Angebot einer automatischen Vektorisierung ein, selbst wenn Tests gute Ergebnisse geliefert haben! Überlassen Sie dem Anbieter die Wahl oder die Kombination der Verfahren und definieren Sie ihm die gewünschten Ergebnisse und deren Qualitätsanforderungen. Auf welche Weise er dann die Ergebnisqualität herstellt, bleibt seinen Möglichkeiten überlassen. Auf diese Weise schließen Sie sich von den (Rest-)Risiken der automatischen Vektorisierung aus. Prüfen Sie immer die durch einen Dienstleister bereit gestellten Ergebnisse der Auftragsarbeit und reklamieren Sie rechtzeitig mangelnde Qualität! Nur so sind Sie sicher, dass Sie für Ihr Geld auch die geforderte Leistung bekommen.

- Stellen Sie bei Datenerfassungsaufträgen die Begrifflichkeit klar. Dadurch räumen Sie früh genug potenzielle Missverständnisse aus! Unterscheiden Sie deutlich die Digitalisierungsart **Bildschirmdigitalisierung** von der (Bildschirm-)**Konstruktion!** Erläutern Sie vorab, was Sie darunter verstehen und stellen Sie die Unterschiede der Verfahren im Hinblick auf Ihren Bedarf heraus. Wie Sie aus dem Kontext leicht erkennen, gibt es zu den Datenerfassungsverfahren keine eindeutigen Begriffsnormen, so dass sonst der Spielraum für Missverständnisse eröffnet wird.

Literatur:

Klemmer, Wilfried: „GIS-Projekte erfolgreich durchführen“
Bernhard Harzer Verlag
Karlsruhe 2004