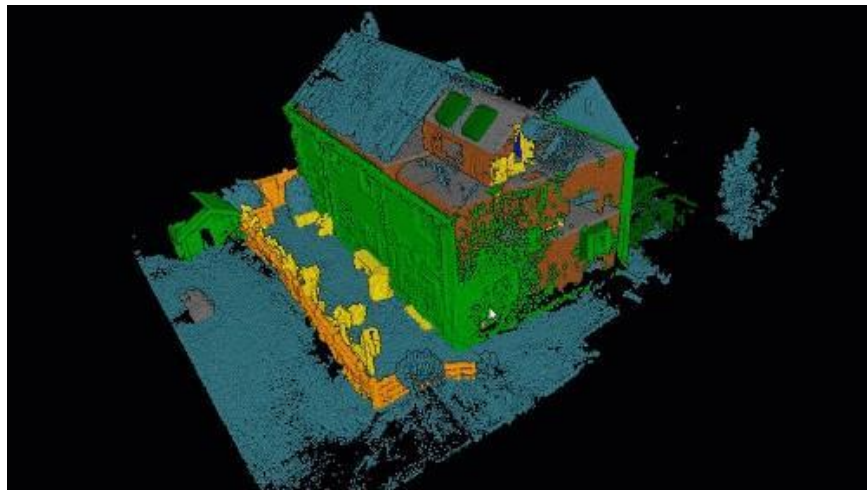




Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Αγρονόμων και Τοπογράφων Μηχανικών
ΔΠΜΣ Γεωπληροφορικής

Μεταπτυχιακή Διπλωματική Εργασία

**Βιβλιογραφική Ανασκόπηση των μεθόδων για την
Σημασιολογική Κατάτμηση Νεφών Σημείων**



Μαρίνα Γ. Τζοβάρα

Επιβλέπων: καθ. Ανδρέας Γεωργόπουλος,

Αθήνα, Φεβρουάριος 2023

Ευχαριστίες

Θα ήθελα να ευχαριστήσω όλους όσους συνέβαλαν στην διεκπεραίωση της παρούσας διπλωματικής εργασίας. Αρχικά, θα ήθελα να ευχαριστήσω τον καθηγητή μου κ. Ανδρέα Γεωργόπουλο για την εμπιστοσύνη που μου έδειξε με την ανάθεση της παρούσας διπλωματικής εργασίας και τη βοήθεια που μου παρείχε.

Στη συνέχεια θα ήθελα να ευχαριστήσω την οικογένεια μου, τον σύντροφό μου και τους φίλους μου για την πολύτιμη στήριξή τους και την ενθάρρυνσή τους αυτή την απαιτητική περίοδο της εκπόνησης της παρούσας διπλωματικής.

*Αθήνα, Φεβρουάριος 2023
Τζοβάρα Μαρίνα*

ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο του Διεπιστημονικού Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών “Γεωπληροφορική” της Σχολής Αγρονόμων Τοπογράφων Μηχανικών – Μηχανικών Γεωπληροφορικής, του Εθνικού Μετσόβιου Πολυτεχνείου. Το αντικείμενό της είναι η βιβλιογραφική ανασκόπηση μεθόδων για τη σημασιολογική κατάτμηση νεφών σημείων.

Η φωτογραμμετρία είναι μια επιστήμη η οποία έχει εξελιχθεί ραγδαία τα τελευταία χρόνια με αποτέλεσμα τα προϊόντα που δημιουργούνται λόγω της ανάπτυξής της να γίνονται όλο και πιο σημαντικά. Ένα από αυτά είναι τα νέφη σημείων, τα οποία αποτελούν σήμερα μια μορφή δεδομένων πολύ χρήσιμη, ταυτόχρονα όμως και πολύ ελκυστική για τη διερεύνηση μεθόδων και τεχνικών ώστε να υπόκεινται σε επεξεργασία πιο εύκολα και πιο αυτόματα. Έτσι, με τα χρόνια δημιουργήθηκε η ανάγκη για την ανάπτυξη μεθόδων για τη σημασιολογική κατάτμηση ενός νέφους σημείων.

Σημασιολογική κατάτμηση ενός νέφους σημείων είναι η διαδικασία κατά την οποία τα σημεία από τα οποία αποτελείται ένα νέφος σημείων χωρίζονται σε κατηγορίες έχοντας κοινά χαρακτηριστικά και μια ιδιότητα. Πριν εφαρμοστούν ευρέως αποτελεσματικές μέθοδοι μάθησης με επίβλεψη στη σημασιολογική κατάτμηση, η κατάτμηση νέφους σημείων χωρίς επίβλεψη ήταν ένα σημαντικό έργο για δεδομένα 2,5D/3D. Η κατάτμηση νέφους σημείων στοχεύει στην ομαδοποίηση σημείων με παρόμοια γεωμετρικά/φασματικά χαρακτηριστικά χωρίς να λαμβάνει υπόψη τις σημασιολογικές πληροφορίες. Επειδή στη ροή εργασίας της σημασιολογικής κατάτμησης, η απλή κατάτμηση μπορεί να χρησιμοποιηθεί ως βήμα προ-τμηματοποίησης, επηρεάζοντας τα τελικά αποτελέσματα, στη παρούσα εργασία θα αναλυθούν τόσο οι τεχνικές κατάτμησης όσο και οι τεχνικές σημασιολογικής κατάτμησης νεφών σημείων.

Επειδή στην ελληνική βιβλιογραφία δεν υπάρχουν πολλές μελέτες που να παρουσιάζουν αυτές τις τεχνικές και να αποσαφηνίζουν τι είναι η κατάτμηση και τι η σημασιολογική κατάτμηση ενός νέφους σημείων, η παρούσα διπλωματική εργασία επιχειρεί να κάνει ακριβώς αυτό.

Λέξεις κλειδιά: νέφος σημείων, κατάτμηση, σημασιολογική κατάτμηση, βαθιά μάθηση.

Αθήνα

Φεβρουάριος 2023

ABSTRACT

The present thesis was prepared in the framework of the Interdisciplinary Postgraduate Programme "Geoinformatics" of the School of Rural, Surveying and Geoinformatics Engineering, of National Technical University of Athens. The subject of this thesis is a literature review of methods for point cloud semantic segmentation.

Photogrammetry is a science that has evolved rapidly in recent years with result that the "products" created due to its development are becoming increasingly important. One of these is point clouds, which are nowadays a very useful but also very attractive form of data for exploring methods and techniques to process them more easily and automatically. Thus, over the years, the need has arisen to explore methods for the point cloud semantic segmentation.

Semantic segmentation of a point cloud is the process by which the points that make up a point cloud are divided into categories having common characteristics and a property. Before widely effective supervised learning methods were widely applied to semantic segmentation, unsupervised point cloud segmentation was an important task for 2.D/3D data. Point cloud segmentation aims to group points with similar geometric/spectral characteristics without taking semantic information into account. Because of the fact that in the semantic segmentation workflow, simple segmentation can be used as a pre-segmentation step, affecting the final results, this paper will analyze both segmentation techniques and point cloud semantic segmentation techniques.

Since there are not many studies in the Greek literature that present these techniques and clarify what is partitioning and what is semantic partitioning of a point cloud, this is what this thesis attempts to do.

Keywords: point cloud, segmentation, semantic segmentation, deep learning.

Athens

February 2023

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ	3
ABSTRACT	4
ΕΙΣΑΓΩΓΗ	7
1. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	8
1.1 Έννοια της Φωτογραμμετρίας	8
1.2 Νέφος Σημείων	10
1.3 Σημασιολογική Κατάτμηση	11
2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ	12
2.1 Τεχνικές Κατάτμησης Νεφών Σημείων	12
2.1.1 Κατάτμηση βάσει των ακμών	12
2.1.2 Κατάτμηση βάσει ανάπτυξης περιοχών	13
2.1.3 Κατάτμηση με βάση την προσαρμογή μοντέλων	15
2.1.4 Κατάτμηση βάσει ομαδοποίησης	18
2.2 Τεχνικές Σημασιολογικής Κατάτμησης Νεφών Σημείων	20
2.2.1 Κλασσική επιβλεπόμενη μηχανική μάθηση	20
2.2.2 Βαθιά Μηχανική Μάθηση	22
2.2.2.1 Σημασιολογική κατάτμηση βασισμένη στη μέθοδο πολλαπλών προβολών (Multiview-based)	23
2.2.2.2 Σημασιολογική κατάτμηση βασισμένη σε ογκομετρικά εικονοστοιχεία (voxel):	24
2.2.2.3 Άμεση επεξεργασία δεδομένων νέφους σημείων (Directly process point cloud data)	26
3. ΕΦΑΡΜΟΓΕΣ ΣΗΜΑΣΙΟΛΟΓΙΚΗΣ ΚΑΤΑΤΜΗΣΗΣ	29
3.1 Πολιτισμική Κληρονομιά.	29
3.1.1 Μεθοδολογία	29
3.1.1.1 Δημιουργία συνόλου δεδομένων	30
3.1.1.2 Προεπεξεργασία	32
3.1.1.3 Σημασιολογική Κατάτμηση PointNet++	33
3.1.1.4 Αξιολόγηση αποτελεσμάτων	35
3.2 Αστικού Τύπου Δεδομένα	36
3.2.1 Μεθοδολογία	37
3.2.1.1 Δεδομένα	37
3.2.1.2 Εξαγωγή χαρακτηριστικών	38
3.2.1.3 Εφαρμογή τεχνικών μηχανικής μάθησης	39

3.2.1.4	Αποτελέσματα	39
4.	ΣΥΜΠΕΡΑΣΜΑΤΑ	41
5.	ΒΙΒΛΙΟΓΡΑΦΙΑ	43

Ευρετήριο Εικόνων

<i>Εικόνα 1:</i>	<i>Παράδειγμα ψευδοεπιπέδου. Δύο καλά εκτιμώμενα επίπεδα υποθέσεων εμφανίζονται με μπλε χρώμα. Με το ίδιο κατώφλι δημιουργείται ένα παραπλανητικό επίπεδο (με πορτοκαλί χρώμα).</i>	17
<i>Εικόνα 2:</i>	<i>Οικογένεια RANSAC με αλγορίθμους κατηγοριοποιημένους ανάλογα με τις επιδόσεις και τις βασικές στρατηγικές τους.</i>	18
<i>Εικόνα 3:</i>	<i>Το πλαίσιο της σημασιολογικής κατάτμησης νεφών σημείων.</i>	20
<i>Εικόνα 4:</i>	<i>Το πλαίσιο της σημασιολογικής κατάτμησης νέφους σημείων.</i>	22
<i>Εικόνα 5:</i>	<i>Το πλαίσιο του νευρωνικού δικτύου πολλαπλών όψεων (MVCNN)</i>	23
<i>Εικόνα 6:</i>	<i>Η ροή εργασιών της μεθόδου SnapNet.</i>	24
<i>Εικόνα 7:</i>	<i>Η αρχιτεκτονική του Voxnet.</i>	25
<i>Εικόνα 8:</i>	<i>Το πλαίσιο SegCloud</i>	26
<i>Εικόνα 9:</i>	<i>Το πλαίσιο του PointNet.</i>	27
<i>Εικόνα 10:</i>	<i>Ροή εργασιών μελέτης των Malinverni κ.ά .</i>	30
<i>Εικόνα 11:</i>	<i>Γενικό σχήμα της ενοποιημένης μεθόδου για τον ορισμό και την ταξινόμηση του αρχιτεκτονικού στοιχείου.</i>	31
<i>Εικόνα 12:</i>	<i>Οπτικοποίηση του ορισμού των αρχιτεκτονικών τάξεων. Στα αριστερά το σαρωμένο αντικείμενο, στα δεξιά κάθε συγκεκριμένη κλάση.</i>	32
<i>Εικόνα 13:</i>	<i>Διαχωρισμός διαφόρων σημασιολογικών κλάσεων με χρώματα</i>	37

ΕΙΣΑΓΩΓΗ

Η σημασιολογική κατάτμηση έχει απασχολήσει αρκετά τους ερευνητές τα τελευταία χρόνια. Οι έρευνες ξεκίνησαν από την δεκαετία του 1970 και στοχεύουν στην κατηγοριοποίηση κάθε εικονοστοιχείου ή σημείου σε διάφορες περιοχές με συγκεκριμένες σημασιολογικές κατηγορίες. Στην αρχή εφαρμόστηκε σε διδιάστατες εικόνες και στην συνέχεια με την εξέλιξη της τεχνολογίας εφαρμόστηκε σε νέφη σημείων.

Κομβικό στάδιο, όμως για την σημασιολογική κατάτμηση είναι η ανάπτυξη τεχνικών για την κατάτμηση των νεφών σημείων, δηλαδή για την ομαδοποίηση σημείων με παρόμοια γεωμετρικά/φασματικά χαρακτηριστικά χωρίς να λαμβάνει υπόψη τις σημασιολογικές πληροφορίες. Γι' αυτό το λόγο στην παρούσα εργασία παρουσιάζονται τόσο οι τεχνικές για την κατάτμηση νεφών σημείων όσο και οι τεχνικές για τη σημασιολογική κατάτμηση νεφών σημείων. Πιο συγκεκριμένα, η διάρθρωση της παρούσας διπλωματικής εργασίας είναι ως εξής:

Στο πρώτο κεφάλαιο γίνεται αναφορά στο θεωρητικό πλαίσιο στο οποίο υπάγεται η σημασιολογική κατάτμηση νεφών σημείων. Αναλύεται η έννοια της φωτογραμμετρίας, τα πλεονεκτήματα και τα μειονεκτήματά της καθώς και η εξέλιξή της ανά τα χρόνια. Έπειτα περιγράφεται η έννοια του νέφους σημείων και τέλος αποσαφηνίζεται τι είναι σημασιολογική κατάτμηση.

Στο δεύτερο κεφάλαιο παρουσιάζονται τόσο οι τεχνικές κατάτμησης όσο και οι τεχνικές σημασιολογικής κατάτμησης.

Στο τρίτο κεφάλαιο παρουσιάζονται δύο μελέτες που διαχωρίζονται βάσει των δεδομένων τους, δηλαδή η μία αφορά σε δεδομένα πολιτιστικής κληρονομιάς και η άλλη σε δεδομένα αστικού τύπου. Είναι δύο παραδείγματα όπου φαίνεται πώς εφαρμόζονται οι τεχνικές σημασιολογικής κατάτμησης σε νέφος σημείων.

Τέλος, στο τέταρτο κεφάλαιο παρουσιάζονται τα συμπεράσματα που εξάγονται από την ανασκόπηση.

1. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Στο παρόν κεφάλαιο, αρχικά, αναλύεται η έννοια της φωτογραμμετρίας. Έπειτα, περιγράφεται η δημιουργία νεφών σημείων και τέλος τι είναι σημασιολογική κατάτμηση και ποιος είναι ο σκοπός της παρούσας εργασίας.

1.1 Έννοια της Φωτογραμμετρίας

Φωτογραμμετρία ονομάζεται “ η τέχνη, επιστήμη και τεχνολογία για την απόκτηση αξιόπιστης πληροφορίας σχετικά με φυσικά αντικείμενα και το περιβάλλον μέσα από διαδικασίες καταγραφής, μέτρησης και ερμηνεία φωτογραφικών εικόνων αλλά και προτύπων ηλεκτρομαγνητικής ακτινοβολίας” (Πατιάς 1991). Ουσιαστικά η φωτογραμμετρία είναι μια τεχνική η οποία καταγράφει με έμμεσο τρόπο τη θέση, το σχήμα και τις διαστάσεις των αντικειμένων διότι δεν πραγματοποιούνται απευθείας μετρήσεις πάνω στα αντικείμενα, αλλά χρησιμοποιεί την μετρητική πληροφορία που συλλέγεται πάνω σε φωτογραφίες. Ωστόσο με την εξέλιξη της τεχνολογίας, η φωτογραμμετρία έχει εξελιχθεί και στην καταγραφή σημάτων από ηλεκτρομαγνητικές (π.χ. laser scanner) και ηχητικές πηγές (π.χ. πολυδесμικοί ηχοβολιστές).

Η έννοια της φωτογραμμετρίας σε μια πιο γενική μορφή, δηλαδή οι τεχνικές μέτρησης πάνω σε εικόνα είναι πιο παλιά και από την ίδια την φωτογραφία. Συγκεκριμένα, το 350 π.Χ. ο Αριστοτέλης αναφέρει την αρχή λειτουργίας του σκοτεινού θαλάμου και θέτει τα θεμέλια της προβολικής γεωμετρίας στην οποία βασίζεται η φωτογραμμετρία. Η ανακάλυψη της φωτογραφίας πραγματοποιήθηκε το 1839 μ.Χ. από τους J.N. Niepce και J.L.M. Daguerre. Τότε ο Arago το ανακοίνωσε στην Γαλλική Ακαδημία και αναφέρει για πρώτη φορά πιθανές εφαρμογές στην απεικόνιση του εδάφους. Ορόσημα της ιστορίας της φωτογραμμετρίας είναι το 1851 διότι τότε κατασκευάστηκε και εφαρμόστηκε η πρώτη μετρητική κάμερα σε τοπογραφικές αποτυπώσεις από τον A. Laussedat ο οποίος θεωρείται και ο πατέρας της φωτογραμμετρίας. Από τότε μέχρι σήμερα, η φωτογραμμετρία εξελίχθηκε ραγδαία (Πατιάς 1991).

Η φωτογραμμετρία έχει μεγάλη σχέση με αντίστοιχες επιστήμες συλλογής πρωτογενών δεδομένων όπως η Τοπογραφία και η Γεωδαισία. Η βασική διαφορά των φωτογραμμετρικών μεθόδων με αυτές που χρησιμοποιούνται στις άλλες επιστήμες είναι ότι οι μετρήσεις γίνονται έμμεσα και πάνω σε κάποιο μέσο και όχι απευθείας στο φυσικό κόσμο. Από την διαφορά αυτή πηγάζουν τα πλεονεκτήματα και τα μειονεκτήματα της Φωτογραμμετρίας από τα οποία τα πιο βασικά είναι τα παρακάτω:

Πλεονεκτήματα

- Ταχύτητα: Με τις φωτογραμμετρικές μεθόδους η παραγωγή χαρτών υλοποιείται πολύ πιο γρήγορα από οποιαδήποτε άλλη μέθοδο που στηρίζεται σε επίγειες μετρήσεις. Αυτό συμβαίνει διότι με τις αεροφωτογραφίες καλύπτεται μια μεγάλη περιοχή π.χ. μια λήψη από ύψος πτήσης 1.5km καλύπτεται μια περιοχή μεγαλύτερη από 5km² , ενώ αντίστοιχη περιοχή θα απαιτούσε πολύ παραπάνω χρόνο για να καλυφθεί από επίγειες μετρήσεις.
- Οικονομία: Το κόστος συνδέεται άμεσα με την ταχύτητα συλλογής δεδομένων. Γι’ αυτό όταν υπάρχει η ανάγκη για μαζική χαρτογράφηση περιοχών

χρησιμοποιούνται φωτογραμμετρικές μέθοδοι καθώς θεωρούνται οι πιο ενδεδειγμένες μέθοδοι ως προς την ταχύτητα και το κόστος.

- Προσπελασιμότητα εδάφους: Είναι εύκολα κατανοητό ότι σε περιοχές δύσβατες ή καθόλου προσβάσιμες προτιμάται η χρήση φωτογραμμετρικών μετρήσεων έναντι των επίγειων
- Συνεχής απεικόνιση: Η υπεροχή της Φωτογραμμετρίας γίνεται εύκολα αντιληπτή από το γεγονός ότι όλη η μετρητική πληροφορία υπάρχει σε μια φωτογραφία. Για παράδειγμα η χαρτογραφική απόδοση της όψης ενός κτιρίου απαιτεί την αποτύπωση μεγάλου αριθμού λεπτομερειών κάτι το οποίο είναι σχεδόν απαγορευτικό να υλοποιηθεί από επίγειες μετρήσεις. Γι' αυτό υλοποιείται με φωτογραμμετρικές μεθόδους όπου όλη η απαιτούμενη πληροφορία υπάρχει στη φωτογραφία.

Μειονεκτήματα

- Πολλές πηγές σφαλμάτων: κατά την συλλογή δεδομένων με φωτογραμμετρικές μεθόδους, οι πηγές σφαλμάτων ποικίλλουν. Για να εξαλειφθούν αυτά τα σφάλματα απαιτούνται πολυπλοκότερες μεθοδολογίες επεξεργασίας δεδομένων ενώ ενέχεται ο κίνδυνος τα τελικά προϊόντα να είναι μερικές φορές μικρότερης ακρίβειας από ό,τι αν είχαν συλλεχθεί με επίγειες μετρήσεις.
- Απαιτήσεις σε σημεία ελέγχου: ορισμένες φορές η φωτογραμμετρική μεθοδολογία βασίζεται σε κάποια ήδη υπάρχουσα πληροφορία απαιτώντας έτσι την ύπαρξη γνωστών σημείων στο έδαφος. Με αυτό τον τρόπο εξαρτάται από το αποτέλεσμα προηγούμενων επίγειων μετρήσεων καθιστώντας της έτσι ένα δευτερογενές στάδιο συλλογής πληροφορίας.

Η εξέλιξη της τεχνολογίας έχει βοηθήσει στη σημαντική μείωση των μειονεκτημάτων των φωτογραμμετρικών μετρήσεων σε σημείο να ανταγωνίζεται σημαντικά πλέον τις επίγειες μετρήσεις. Ένα σημαντικό βήμα της εξέλιξης της φωτογραμμετρίας είναι η δημιουργία νεφών σημείων και αυτό αποδεικνύεται από το γεγονός ότι πλέον έχουν πολύ σημαντικό ρόλο σε διάφορους επιστημονικούς τομείς.

1.2 Νέφος Σημείων

Νέφος σημείων είναι μια τρισδιάστατη απεικόνιση η οποία αποτελείται από ένα σύνολο σημείων τα οποία έχουν συγκεκριμένες συντεταγμένες (x,y,z) και έχουν την δυνατότητα να λαμβάνουν χρωματική πληροφορία του αντικειμένου που αντιστοιχεί στο κάθε σημείο. Τα νέφη σημείων παράγονται γενικά από τρισδιάστατους σαρωτές ή από λογισμικό φωτογραμμετρίας, τα οποία μετρούν πολλά σημεία στις εξωτερικές επιφάνειες των αντικειμένων γύρω τους.

Οι τρισδιάστατοι σαρωτές είναι συσκευές οι οποίες καταγράφουν ένα πραγματικό αντικείμενο για να συλλέξουν από αυτό στοιχεία που αφορούν στη μορφή του αντικειμένου. Κατά τη διάρκεια της σάρωσης χιλιάδες μεμονωμένες τρισδιάστατες μετρήσεις εμφανίζονται δημιουργώντας σε πραγματικό χρόνο μια τρισδιάστατη απεικόνιση της λήψης. Από τις πολικές συντεταγμένες υπολογίζονται στη συνέχεια αυτόματα οι καρτεσιανές συντεταγμένες (x, y, z) των σημείων του προς αποτύπωση αντικειμένου στο τρισσορθογώνιο σύστημα του σαρωτή. Έτσι, σε κάθε σημείο αντιστοιχεί η πληροφορία (x, y, z) , με συνέπεια να ορίζεται ακριβώς η θέση του σημείου. Με αυτό τον τρόπο δημιουργούνται τα νέφη σημείων από τους σαρωτές.

Νέφη σημείων μπορούν να παραχθούν και με έμμεσο τρόπο μέσω εικόνων. Αποκτώντας στερεοσκοπικές εικόνες ή εικόνες πολλαπλών όψεων μέσω ηλεκτροοπτικών συστημάτων π.χ. κάμερες, υπολογίζονται πληροφορίες για κάθε τρισδιάστατο σημείο σύμφωνα με τις αρχές της θεωρίας της φωτογραμμετρίας ή της όρασης υπολογιστών. Η εξέλιξη της τεχνολογίας έχει βοηθήσει στο να υπάρχουν πλέον πιο αυτόματοι τρόποι για την παραγωγή νέφους σημείων από εικόνες. Η πυκνή συνταύτιση (dense matching), η στερεοσκοπική απεικόνιση πολλαπλών όψεων (Multiple View Stereovision) και η δομή από κίνηση (Structure for Motion) άλλαξαν το νέφος σημείων που προέκυπτε από την εικόνα και άνοιξαν την εποχή της πολυεικονικής στερεοσκοπικής απεικόνισης. Η δομή από κίνηση (SfM) μπορεί να εκτιμήσει αυτόματα τις θέσεις και τους προσανατολισμούς της κάμερας έχοντας ως αποτέλεσμα την δυνατότητα επεξεργασίας εικόνων πολλαπλών προβολών ταυτόχρονα, ενώ οι αλγόριθμοι πυκνής συνταύτισης (DM) και η πολυεικονική στερεοσκοπική απεικόνιση (MVS) παρέχουν τη δυνατότητα παραγωγής μεγάλου όγκου νεφών σημείων. Τα τελευταία χρόνια, τα πλήρη τρισδιάστατα πυκνά νέφη σημείων σε κλίμακα πόλης μπορούν να αποκτηθούν εύκολα μέσω μιας τεχνικής πλάγιας φωτογράφισης που βασίζεται σε δομή από κίνηση (SfM) και πολυεικονική στερεοσκοπική απεικόνιση (MVS) (Xie κ.ά. 2020).

Το νέφος σημείων έχει γίνει πλέον πόλος έλξης των ερευνητών για τα τρισδιάστατα δεδομένα και συνεισφέρει πολύ στην πλοήγηση σε εσωτερικούς χώρους, στην μη επανδρωμένη οδήγηση, στην ανάλυση της αστικής μορφολογίας, στην προστασία της ψηφιακής πολιτιστικής κληρονομιάς και σε πολλές άλλες εφαρμογές.

1.3 Σημασιολογική Κατάτμηση

Σήμερα, τα τρισδιάστατα μοντέλα και τα νέφη σημείων χρησιμοποιούνται σε διάφορα πεδία έχοντας γίνει αρκετά γνωστά και χρήσιμα. Παρ' όλο που γίνονται ολοένα και πιο εύκολα διαθέσιμα, υπάρχει ακόμα η ανάγκη για μεθόδους ιδιαίτερα αυτοματοποιημένες, ώστε στα τρισδιάστατα δεδομένα να παρέχονται σημαντικά χαρακτηριστικά που χαρακτηρίζουν και προσδίδουν σημασία στα αντικείμενα που αναπαρίστανται σε 3D. Η κατάτμηση είναι η διαδικασία κατά την οποία ομαδοποιούνται τα νέφη σημείων σε πολύπλοκες ομοιογενείς περιοχές με ίδιες ιδιότητες.

Η σημασιολογική κατάτμηση είναι μια από τις πιο σημαντικές ερευνητικές τεχνολογίες, η οποία πρωτοεμφανίστηκε τη δεκαετία του 1970 και στοχεύει στην κατηγοριοποίηση κάθε εικονοστοιχείου ή σημείου σε διάφορες περιοχές με συγκεκριμένες σημασιολογικές κατηγορίες. Στην αρχή εφαρμόστηκε σε δισδιάστατες εικόνες έχοντας σημειώσει μεγάλη εξέλιξη τα τελευταία χρόνια. Ωστόσο, λόγω των περιορισμών των δισδιάστατων δεδομένων, η απόδοση της κατάτμησης δεν είναι ικανοποιητική. Γι' αυτόν το λόγο οι ερευνητές στρέφουν σταδιακά την προσοχή τους σε τρισδιάστατα δεδομένα όπως τρισδιάστατα πλέγματα ογκομετρικών στοιχείων (voxel) ή τρισδιάστατα νέφη σημείων (Zhang κ.ά. 2019).

Η σημασιολογική κατάτμηση νεφών σημείων είναι η τρισδιάστατη μορφή σημασιολογικής κατάτμησης στην οποία χρησιμοποιούνται ακανόνιστα κατανεμημένα σημεία στον τρισδιάστατο χώρο αντί για κανονικοποιημένα κατανεμημένα εικονοστοιχεία σε μια δισδιάστατη εικόνα. Η έρευνα σχετικά με την σημασιολογική κατάτμηση νεφών σημείων έχει μακρά παράδοση που περιλαμβάνει διαφορετικά πεδία και ορίζει διαφορετικές έννοιες για παρόμοια καθήκοντα δηλαδή χρησιμοποιείται διαφορετικός όρος για την ίδια έννοια σε διαφορετικά επιστημονικά πεδία. Για παράδειγμα, στην όραση υπολογιστών, ιδίως σε πρόσφατες εφαρμογές βαθιάς μάθησης χρησιμοποιείται ευρέως ο όρος σημασιολογική κατάτμηση νεφών σημείων (Qi κ.ά. 2016)-(Wang κ.ά. 2019). Ωστόσο, στη φωτογραμμετρία και την τηλεπισκόπηση, η σημασιολογική κατάτμηση νεφών σημείων ονομάζεται συνήθως "ταξινόμηση νέφους σημείων" (Zhang κ.ά. 2013)-(Wang κ.ά. 2015). Στην παρούσα εργασία, σημασιολογική κατάτμηση νέφους σημείων θεωρείται η συσχέτιση κάθε σημείου ενός νέφους σημείων με μια σημασιολογική ετικέτα έτσι ώστε να συμβαδίζει η παρούσα βιβλιογραφική ανασκόπηση με τις σύγχρονες μεθόδους οι οποίες είναι οι τεχνικές βαθιάς μάθησης.

Η κατάτμηση νέφους σημείων στοχεύει στην ομαδοποίηση σημείων με παρόμοια γεωμετρικά/φασματικά χαρακτηριστικά χωρίς να λαμβάνει υπόψη τις σημασιολογικές πληροφορίες. Στη ροή εργασίας της σημασιολογικής κατάτμησης, η απλή κατάτμηση μπορεί να χρησιμοποιηθεί ως βήμα προ-τμηματοποίησης, επηρεάζοντας τα τελικά αποτελέσματα (Xie κ.ά. 2020). Γι' αυτόν το λόγο στην παρούσα εργασία θα αναλυθούν τόσο οι τεχνικές κατάτμησης όσο και οι τεχνικές σημασιολογικής κατάτμησης νεφών σημείων.

2. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

Όπως ήδη έχει αναφερθεί, σημασιολογική κατάτμηση νεφών σημείων είναι η συσχέτιση κάθε σημείου ενός νέφους σημείων με μια σημασιολογική ιδιότητα. Πιο απλά, είναι η διαδικασία κατά την οποία τα σημεία από τα οποία αποτελείται ένα νέφος σημείων χωρίζονται σε κατηγορίες που παρουσιάζουν κοινά χαρακτηριστικά και μια ιδιότητα. Στον παρόν κεφάλαιο αναλύονται αρχικά οι τεχνικές κατάτμησης των νεφών σημείων, οι οποίες αποτελούν κομβικό στάδιο για την εξέλιξη της σημασιολογικής κατάτμησης και έπειτα παρουσιάζονται οι σύγχρονες τεχνικές σημασιολογικής κατάτμησης.

2.1 Τεχνικές Κατάτμησης Νεφών Σημείων

Οι αλγόριθμοι κατάτμησης νεφών σημείων βασίζονται κυρίως σε αυστηρά χαρακτηριστικά από γεωμετρικούς περιορισμούς και στατιστικούς κανόνες. Η κύρια διαδικασία της κατάτμησης νεφών σημείων αποσκοπεί στην ομαδοποίηση τρισδιάστατων σημείων καθένα από τα οποία μπορεί να ανήκει σε μία ομάδα. Οι ομάδες αυτές αντιστοιχούν σε συγκεκριμένες δομές ή αντικείμενα σε μια σκηνή (π.χ στέγες σπιτιών, βλάστηση κλπ). Δεδομένου ότι δεν απαιτείται καμία εποπτευόμενη προηγούμενη γνώση σε μια τέτοια διαδικασία κατάτμησης, τα αποτελέσματα που παράγονται δεν έχουν ισχυρή σημασιολογική πληροφορία. Αυτές οι προσεγγίσεις μπορούν να κατηγοριοποιηθούν σε τέσσερις μεγάλες ομάδες: με βάση τις ακμές, την ανάπτυξη περιοχών, την προσαρμογή μοντέλων και την ομαδοποίηση (Xie κ.ά. 2020).

2.1.1 Κατάτμηση βάσει των ακμών

Το σχήμα των αντικειμένων περιγράφεται από τις ακμές συνεπώς η κατάτμηση των νεφών σημείων μπορεί να επιλυθεί βρίσκοντας τα σημεία που βρίσκονται κοντά στις περιοχές των ακμών. Η αρχή των μεθόδων που βασίζονται στις ακμές είναι να εντοπίζουν τα σημεία που έχουν μια ταχεία αλλαγή στην ένταση (Nguyen και Le 2013), κάτι που είναι παρόμοιο με ορισμένες προσεγγίσεις τμηματοποίησης εικόνων 2D.

Η κατάτμηση βάσει ακμών έχει δύο κύρια στάδια. Το πρώτο είναι η ανίχνευση ακμών, όπου εξάγονται τα όρια των διαφόρων περιοχών και το δεύτερο η ομαδοποίηση των σημείων που βρίσκονταν μέσα στα όρια δημιουργώντας τα τελικά τμήματα. Οι Bhanu κ.ά. το 1986 σχεδίασαν έναν αλγόριθμο βασισμένο στην κλίση για την ανίχνευση ακμών, προσαρμόζοντας τρισδιάστατες γραμμές σε ένα σύνολο σημείων και ανιχνεύοντας αλλαγές στην κατεύθυνση των μοναδιαίων κάθετων διανυσμάτων στην επιφάνεια. Οι Sappa and Denvy το 2001 εξήγαγαν στενά περιγράμματα από ένα δυαδικό χάρτη ακμών για γρήγορη κατάτμηση. Τέλος, οι Wani and Arbania το 2003 εισήγαγαν έναν παράλληλο αλγόριθμο κατάτμησης με βάση τις ακμές που εξάγει τρεις τύπους ακμών. Ένας μηχανισμός βελτιστοποίησης του αλγορίθμου, με την ονομασία reconfigurable multiRing network, εφαρμόστηκε σε αυτόν τον αλγόριθμο για να μειώσει τον χρόνο εκτέλεσής του (Xie κ.ά. 2020).

Οι αλγόριθμοι κατάτμησης με βάση τις ακμές επιτυγχάνουν μια γρήγορη κατάτμηση νεφών σημείων χάρη στην απλότητά τους. Ωστόσο τα καλά αποτελέσματά τους μπορεί να επιτευχθούν μόνο όταν παρέχονται απλές σκηνές με ιδανικά σημεία (π.χ. χαμηλός θόρυβος, ομοιόμορφη πυκνότητα). Τα νέφη σημείων συνήθως δεν έχουν αυτά τα χαρακτηριστικά, συνεπώς τέτοιες μέθοδοι δεν θα έχουν τα επιθυμητά αποτελέσματα.

Γι' αυτό εξ άλλου προτιμάται να εφαρμόζονται σε εικόνες παρά σε τρισδιάστατα σημεία. Συνεπώς, η προσέγγιση αυτή σπάνια εφαρμόζεται σήμερα για πυκνά ή/και μεγάλης έκτασης δεδομένα νέφους σημείων. Εξ άλλου στο τρισδιάστατο χώρο τέτοιες μέθοδοι συχνά παρέχουν ασύνδετες ακμές, οι οποίες δεν μπορούν να χρησιμοποιηθούν για τον άμεσο εντοπισμό κλειστών τμημάτων χωρίς διαδικασία συμπλήρωσης ή ερμηνείας (Xie κ.ά. 2020, Castillo κ.ά. 2013, Grilli κ.ά. 2017).

2.1.2 Κατάτμηση βάσει ανάπτυξης περιοχών

Η ανάπτυξη περιοχών (region growing) είναι μια κλασσική μέθοδος κατάτμησης νεφών σημείων η οποία χρησιμοποιείται ευρέως ακόμα και σήμερα. Χρησιμοποιεί κάποια κριτήρια συνδυάζοντας στοιχεία μεταξύ δύο σημείων ή δύο μονάδων περιοχών προκειμένου να μετρήσει τις ομοιότητες μεταξύ των εικονοστοιχείων (2D), σημείων (3D) ή ογκομετρικών εικονοστοιχείων (voxels) (3D) και τα συγχωνεύει αν είναι χωρικά κοντά και έχουν όμοιες επιφανειακές ιδιότητες. Οι Besl και Jain εισήγαγαν έναν αρχικό αλγόριθμο δύο βημάτων. Το πρώτο βήμα είναι η χονδροειδής κατάτμηση κατά την οποία επιλέγονται αρχικά εικονοστοιχεία (seed pixels) βάσει της μέσης τιμής, την γκαουσιανή κατανομή κάθε σημείου και το πρόσημό του. Το δεύτερο βήμα είναι η ανάπτυξη περιοχών κατά την οποία χρησιμοποιείται μια διαδραστική ανάπτυξη περιοχών για την βελτίωση του αποτελέσματος του πρώτου βήματος με βάση προκαθορισμένα κριτήρια όπως η εγγύτητα των σημείων και η επιπεδότητα της επιφάνειας (Xie κ.ά. 2020).

Αρχικά, η μέθοδος αυτή χρησιμοποιήθηκε κυρίως στη δισδιάστατη κατάτμηση καθώς στο πρώιμο στάδιο της έρευνας της κατάτμησης νεφών σημείων τα περισσότερα νέφη σημείων ήταν στην πραγματικότητα 2,5D δεδομένα από αερομεταφερόμενα LiDAR, στα οποία μόνο ένα στρώμα έχει προβολή στη διεύθυνση z, το γενικό βήμα προεπεξεργασίας ήταν να μετατραπούν τα σημεία από τον τρισδιάστατο χώρο σε ένα δισδιάστατο κανονικοποιημένο πλαίσιο. Η μέθοδος της ανάπτυξης περιοχών υιοθετήθηκε απ' ευθείας στον τρισδιάστατο χώρο καθώς γινόταν όλο και πιο εύκολα διαθέσιμα τα πραγματικά τρισδιάστατα νέφη σημείων.

Όπως και στις δισδιάστατες περιπτώσεις, έτσι και στις τρισδιάστατες, η μέθοδος αυτή αποτελείται από δύο βήματα. Το πρώτο είναι η επιλογή των αρχικών σημείων (seed points) ή των αρχικών μονάδων (seed units) και το δεύτερο είναι η ανάπτυξη των περιοχών βασιζόμενη σε κάποιες αρχές. Αυτές οι αρχές είναι οι κρίσιμοι παράγοντες που πρέπει να ληφθούν υπόψη για την δημιουργία ενός αλγορίθμου κατάτμησης ανάπτυξης περιοχών και οι οποίοι είναι:

- Τα κριτήρια (π.χ ίδιες μετρήσεις)
- Μονάδα ανάπτυξης
- Επιλογή αρχικού σημείου (seed point)

Τα κριτήρια

Για τον παράγοντα κριτήρια, συνήθως χρησιμοποιούνται γεωμετρικά χαρακτηριστικά (π.χ. ευκλείδεια απόσταση ή κάθετο διάνυσμα). Για παράδειγμα, οι Ning κ.ά. (Ning κ.ά. 2009) χρησιμοποίησαν το κάθετο διάνυσμα ως κριτήριο, έτσι ώστε να δημιουργούνται περιοχές με σημεία τα οποία ανήκουν στο ίδιο επίπεδο. Οι Tovari et al. (Tovari και Pfeifer 2005) εφάρμοσαν τα κάθετα διανύσματα, την απόσταση των γειτονικών σημείων από το επίπεδο προσαρμογής και την απόσταση μεταξύ του τρέχοντος σημείου και των υποψήφιων σημείων ως κριτήρια για τη συγχώνευση ενός σημείου σε μια αρχική περιοχή που επιλέχθηκε τυχαία από το σύνολο δεδομένων μετά από χειροκίνητο φιλτράρισμα περιοχών κοντά σε ακμές. Οι Dong et al. (Dong κ.ά. 2018) επέλεξαν τα κάθετα διανύσματα και την απόσταση μεταξύ δύο μονάδων.

Μονάδα ανάπτυξης

Για το παράγοντα της μονάδας ανάπτυξης υπάρχουν συνήθως τρεις στρατηγικές. Αυτές είναι τα μεμονωμένα σημεία, οι μονάδες περιοχής, όπως για παράδειγμα τα ογκομετρικά στοιχεία σε πλέγματα (voxel grid) και οι υβριδικές μονάδες. Στην αρχή, η κύρια προσέγγιση ήταν η επιλογή μεμονωμένων σημείων ως μονάδων περιοχής. Αυτό αποδείχθηκε όμως αρκετά χρονοβόρο για μεγάλα νέφη σημείων τα οποία αποτελούνται από χιλιάδες σημεία. Για να μειωθεί ο όγκος επεξεργασίας του ακατέργαστου νέφους σημείων, κατ' επέκταση και ο χρόνος αλλά και να βελτιωθεί η αποτελεσματικότητα των υπολογισμών, διαπιστώθηκε ότι η μονάδα περιοχής είναι μια καλή εναλλακτική λύση. Αυτό συμβαίνει διότι σε ένα νέφος σημείων ο αριθμός ογκομετρικών στοιχείων από τα οποία αποτελείται η περιοχή ως μονάδα ανάπτυξης είναι μικρότερος από τον αριθμό των σημείων από τον οποίο αποτελείται το νέφος σημείων. Με γνώμονα αυτή τη στρατηγική, οι Deschaud κ. ά. (Deschaud και Goulette 2015) παρουσίασαν έναν αλγόριθμο ανάπτυξης περιοχών με βάση τα voxel για τη βελτίωση της αποδοτικότητας αντικαθιστώντας τα σημεία με voxels κατά τη διαδικασία ανάπτυξης περιοχών. Οι Vo κ. ά. (Vo κ.ά. 2015) πρότειναν έναν προσαρμοστικό αλγόριθμο ανάπτυξης περιοχών με βάση την οκτάδα για γρήγορη κατάτμηση επιφανειών, ομαδοποιώντας σταδιακά παρακείμενα voxels με παρόμοιο σημαντικά χαρακτηριστικά. Ως ισορροπία μεταξύ ακρίβειας και αποτελεσματικότητας, προτάθηκαν και δοκιμάστηκαν επίσης υβριδικές μονάδες από διάφορες μελέτες. Για παράδειγμα, οι Xiao κ.ά. (Xiao κ.ά. 2013) συνδύασαν μεμονωμένα σημεία με υποπαράθυρα ως μονάδες ανάπτυξης για τον εντοπισμό επιπέδων. Οι Dong κ. ά. (Dong κ.ά. 2018) χρησιμοποίησαν έναν υβριδικό αλγόριθμο ανάπτυξης περιοχών, βασισμένο σε μονάδες τόσο μεμονωμένων σημείων όσο και supervoxels, για να πραγματοποιήσουν χονδροειδή τμηματοποίηση πριν από τη συνολική βελτιστοποίηση της ενέργειας (Xie κ.ά. 2020).

Αρχικά Σημεία(Seed points)

Πολλοί αλγόριθμοι ανάπτυξης περιοχών στοχεύουν στην κατάτμηση επιπέδου, γι' αυτό το λόγο μια συνήθης πρακτική για την επιλογή αρχικού σημείου (seed points) είναι να σχεδιάζεται πρώτα ένα επίπεδο προσαρμογής για ένα συγκεκριμένο σημείο και των γειτονικών του σημείων και στη συνέχεια να επιλέγεται το σημείο με το ελάχιστο υπόλοιπο στο επίπεδο προσαρμογής ως αρχικό σημείο. Το υπόλοιπο εκτιμάται συνήθως από την απόσταση μεταξύ ενός σημείου και του επιπέδου προσαρμογής του

(Rabbani κ.ά. 2006) -(Ning κ.ά. 2009) ή από την καμπυλότητα του σημείου (Dong κ.ά. 2018).

2.1.3 Κατάτμηση με βάση την προσαρμογή μοντέλων

Η βασική ιδέα της προσαρμογής μοντέλου είναι η αντιστοίχιση των νεφών σημείων με διάφορα γεωμετρικά σχήματα. Γι' αυτόν το λόγο, η μέθοδος αυτή συνήθως θεωρείται μέθοδος ανίχνευσης ή εξαγωγής σχημάτων. Οι πιο ευρέως χρησιμοποιούμενες μέθοδοι προσαρμογής μοντέλων βασίζονται σε δυο κλασικούς αλγόριθμους, τον μετασχηματισμό Hough (HT) και τον RANSAC.

Μετασχηματισμός Hough (HT)

Ο μετασχηματισμός Hough είναι μια κλασική τεχνική ανίχνευσης χαρακτηριστικών στη ψηφιακή επεξεργασία εικόνας. Στην αρχή παρουσιάστηκε για την ανίχνευση γραμμών σε δισδιάστατες εικόνες έχοντας τρία βασικά βήματα. Το πρώτο βήμα είναι η απεικόνιση κάθε δείγματος (π.χ. εικονοστοιχεία σε δισδιάστατες εικόνες και σημεία σε νέφη σημείων) του αρχικού χώρου σε διακριτοποιημένες παραμέτρους. Το δεύτερο βήμα είναι η δημιουργία πίνακα συσσώρευσης με τις παραμέτρους και για κάθε δείγμα εισόδου η επιλογή του βασικού γεωμετρικού στοιχείου του οποίου αποτελούν ακραίες τιμές στο πίνακα των παραμέτρων. Το τρίτο και τελευταίο βήμα είναι η επιλογή του κελιού με την τοπική μέγιστη βαθμολογία, του οποίου οι συντεταγμένες παραμέτρων χρησιμοποιούνται για την αναπαράσταση ενός γεωμετρικού τμήματος στον αρχικό χώρο. Η πιο βασική εκδοχή του μετασχηματισμού Hough είναι ο Γενικευμένος Μετασχηματισμός Hough (Generalized Hough Transform, GHT), που ονομάζεται επίσης Τυποποιημένος Μετασχηματισμός Hough (Standard Hough Transform, SHT) (Duda και Hart 1972). Ο γενικευμένος μετασχηματισμός Hough χρησιμοποιεί μια παραμετροποίηση γωνίας-ακτίνας, προκειμένου να αποφευχθεί το πρόβλημα της άπειρης κλίσης και να απλοποιηθεί ο υπολογισμός. Είναι βασισμένος στον τύπο:

$$\rho = x \cos(\theta) + y \sin(\theta)$$

όπου x και y είναι οι συντεταγμένες εικόνας ενός αντίστοιχου εικονοστοιχείου δείγματος, ρ είναι η απόσταση μεταξύ της αρχής και της γραμμής που διέρχεται από το αντίστοιχο εικονοστοιχείο και θ είναι η γωνία μεταξύ της καθέτου της προαναφερθείσας γραμμής και του άξονα x . Η παραμετροποίηση γωνίας-ακτίνας μπορεί επίσης να επεκταθεί στον τρισδιάστατο χώρο, και έτσι μπορεί να χρησιμοποιηθεί στην ανίχνευση τρισδιάστατων χαρακτηριστικών και στην κατάτμηση κανονικών γεωμετρικών δομών. Σε σύγκριση με τη δισδιάστατη μορφή, στον τρισδιάστατο χώρο, υπάρχει μία ακόμη παράμετρος γωνίας, η φ :

$$\rho = x \cos(\theta) \sin(\varphi) + y \sin(\theta) \sin(\varphi) + z \cos(\varphi)$$

όπου x , y και z είναι οι αντίστοιχες συντεταγμένες ενός τρισδιάστατου δείγματος (π. χ. ένα συγκεκριμένο σημείο από ολόκληρο το νέφος σημείων) και θ και φ είναι οι πολικές συντεταγμένες του κάθετου διανύσματος του επιπέδου, το οποίο περιλαμβάνει το τρισδιάστατο δείγμα.

Ένα από τα σημαντικότερα μειονεκτήματα του γενικευμένου μετασχηματισμού Hough είναι η έλλειψη ορίων στον χώρο των παραμέτρων, η οποία οδηγεί σε υψηλή κατανάλωση μνήμης και μεγάλο χρόνο υπολογισμού. Ως εκ τούτου, έχουν διεξαχθεί ορισμένες μελέτες για τη βελτίωση της απόδοσης του Μετασχηματισμού Hough με τη μείωση του κόστους της διαδικασίας ψηφοφορίας. Τέτοιοι αλγόριθμοι είναι ο Πιθανοτικός Μετασχηματισμός Hough (Probabilistic Hough Transform, PHT), ο Προσαρμοστικός Πιθανοτικός Μετασχηματισμός Hough (Adaptive Probabilistic Hough Transform, APHT), ο Προοδευτικός Πιθανοτικός Μετασχηματισμός Hough (Progressive Probabilistic Hough Transform, PPHT), ο Τυχαίος Μετασχηματισμός Hough (Randomized Hough Transform, RHT) και ο Μετασχηματισμός Hough με βάση τον πυρήνα (Kernel-based Hough Transform, KHT). Εκτός από το υπολογιστικό κόστος, η επιλογή της κατάλληλης αναπαράστασης του συσσωρευτή είναι επίσης ένας τρόπος βελτιστοποίησης της απόδοσης του μετασχηματισμού Hough (Xie κ.ά. 2020).

RANSAC

Η τεχνική RANSAC είναι η άλλη δημοφιλής μέθοδος προσαρμογής μοντέλων. Έχουν δημοσιευτεί αρκετές μελέτες σχετικά με τις γενικές μεθόδους που βασίζονται στην RANSAC. Ο αλγόριθμος που βασίζεται στη συναίνεση τυχαίων δειγμάτων έχει δύο κύριες φάσεις. Η πρώτη φάση είναι η δημιουργία μιας υπόθεσης από τυχαία δείγματα (παραγωγή υποθέσεων) και η δεύτερη φάση είναι η επαλήθευση της υπόθεσης με τα δεδομένα (αξιολόγηση υποθέσεων/επιβεβαίωση μοντέλου) (Choi κ.ά. 2009)-(Raguram κ.ά. 2008). Πριν από τη πρώτη φάση, όπως και στην περίπτωση των μεθόδων που βασίζονται στο μετασχηματισμό Hough, τα μοντέλα πρέπει να οριστούν ή να επιλεγούν χειροκίνητα. Συνήθως στην κατάτμηση των νεφών σημείων και ανάλογα με τη δομή των τρισδιάστατων σκηνών, τα μοντέλα που επιλέγονται είναι επίπεδα, σφαίρες ή άλλα γεωμετρικά πρωτότυπα που μπορούν να αναπαρασταθούν με αλγεβρικούς τύπους.

Κατά τη δημιουργία υποθέσεων, η τεχνική RANSAC επιλέγει τυχαία N σημεία δειγματοληψίας και εκτιμά ένα σύνολο παραμέτρων του μοντέλου χρησιμοποιώντας αυτά τα σημεία δειγματοληψίας. Για παράδειγμα, στην κατάτμηση νεφών σημείων, αν το δεδομένο μοντέλο είναι ένα επίπεδο, τότε $N = 3$, αφού 3 μη-συμμετρικά σημεία καθορίζουν ένα επίπεδο. Το επίπεδο μοντέλο μπορεί να αναπαρασταθεί με:

$$aX + bY + cZ + d = 0$$

όπου $[a, b, c, d]^T$ είναι το σύνολο παραμέτρων προς εκτίμηση.

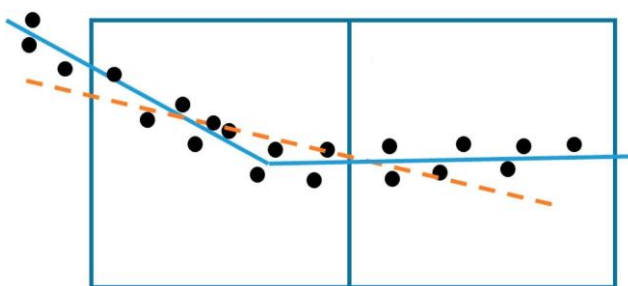
Κατά την αξιολόγηση υποθέσεων, η τεχνική RANSAC επιλέγει την πιο πιθανή υπόθεση από όλα τα εκτιμώμενα σύνολα παραμέτρων. Η τεχνική RANSAC χρησιμοποιεί την παρακάτω εξίσωση για την επίλυση του προβλήματος επιλογής, το οποίο θεωρείται πρόβλημα βελτιστοποίησης (Choi κ.ά. 2009)-(Xie κ.ά. 2020):

$$\hat{M} = \arg \min_M \left\{ \sum_{d \in D} \text{Loss}(\text{Err}(d; M)) \right\}$$

όπου D είναι τα δεδομένα, $Loss$ αντιπροσωπεύει μια συνάρτηση απωλειών και Err είναι μια συνάρτηση σφάλματος, όπως η γεωμετρική απόσταση.

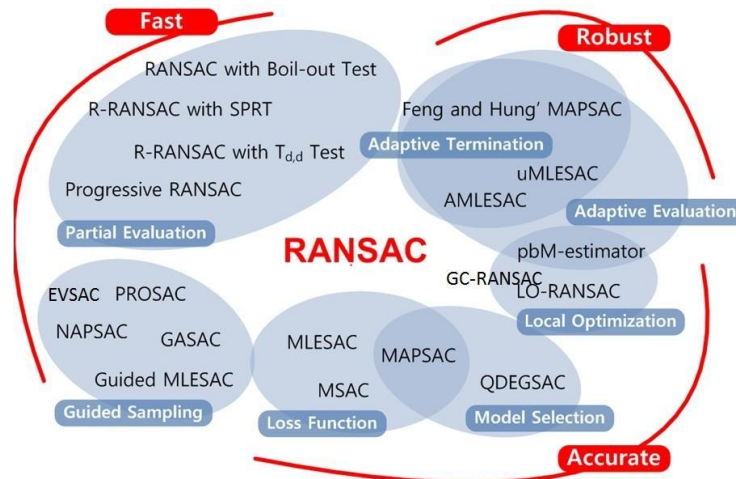
Ως πλεονέκτημα της τυχαίας δειγματοληψίας, οι αλγόριθμοι που βασίζονται στην τεχνική RANSAC δεν απαιτούν πολύπλοκη βελτιστοποίηση ή υψηλούς πόρους μνήμης. Σε σύγκριση με τις μεθόδους του μετασχηματισμού Hough, δύο βασικά πλεονεκτήματα της RANSAC στην τρισδιάστατη κατάτμηση νεφών σημείων είναι η αποδοτικότητα και το ποσοστό των επιτυχώς εντοπισμένων. Επιπλέον, οι αλγόριθμοι RANSAC έχουν την ικανότητα να επεξεργάζονται δεδομένα με μεγάλη ποσότητα θορύβου, ακόμη και ακραίες τιμές. Για την κατάτμηση των νεφών σημείων, όπως ο μετασχηματισμός Hough και η ανάπτυξη περιοχών, ομοίως και η τεχνική RANSAC χρησιμοποιείται ευρέως στην κατάτμηση επιπέδων, όπως οι προσόψεις κτιρίων, οι στέγες κτιρίων και οι σκηνές εσωτερικών χώρων. Σε ορισμένους τομείς υπάρχει ζήτηση για την τμηματοποίηση πιο σύνθετων δομών από τα επίπεδα. Γι' αυτό το λόγο δημιουργήθηκε ένα αυτόματο πλαίσιο αλγορίθμου με βάση την τεχνική RANSAC για την ανίχνευση βασικών γεωμετρικών σχημάτων σε μη οργανωμένα νέφη σημείων. Τα σχήματα αυτά περιλαμβάνουν όχι μόνο επίπεδα, αλλά και σφαίρες, κυλίνδρους και κώνους. Αλγόριθμοι κατάτμησης νεφών σημείων με βάση τη τεχνική RANSAC χρησιμοποιήθηκαν επίσης για κυλινδρικά αντικείμενα (Xie κ.ά. 2020).

Ο RANSAC είναι ένας μη ντετερμινιστικός αλγόριθμος και, ως εκ τούτου, το κύριο μειονέκτημά του είναι η ψεύτικη επιφάνεια: υπάρχει η πιθανότητα τα μοντέλα που ανιχνεύονται από τον αλγόριθμο που βασίζεται στην τεχνική RANSAC να μην υπάρχουν στην πραγματικότητα (Εικόνα 1). Για να ξεπεραστεί η αρνητική επίδραση της τεχνικής RANSAC στη κατάτμηση του νέφους σημείων, δημιουργήθηκε μια συνάρτηση επιλογής κατωφλίου για τη βελτίωση της ποιότητας τμηματοποίησης, στην οποία λαμβάνεται υπόψη τόσο η απόσταση σημείου-επίπεδου όσο και η συνέπεια μεταξύ των κάθετων διανυσμάτων. Μια άλλη βελτιωμένη μέθοδος RANSAC η οποία αποσκοπεί και αυτή στην αποφυγή του προβλήματος της ψευδοεπιφάνειας βασίζεται σε κύτταρα NDT.



Εικόνα 1: Παράδειγμα ψευδοεπιπέδου(Xie κ.ά. 2020). Δύο καλά εκτιμώμενα επίπεδα υποθέσεων εμφανίζονται με μπλε χρώμα. Με το ίδιο κατώφλι δημιουργείται ένα παραπλανητικό επίπεδο (με πορτοκαλί χρώμα).

Όπως και με το μετασχηματισμό Hough, τις τελευταίες δεκαετίες έχουν εμφανιστεί πολλοί βελτιωμένοι αλγόριθμοι που βασίζονται στη τεχνική RANSAC για να βελτιώσουν περαιτέρω την αποδοτικότητα και την ακρίβεια της. Οι προσεγγίσεις αυτές έχουν κατηγοριοποιηθεί ανάλογα με τους ερευνητικούς τους στόχους και παρουσιάζονται στην εικόνα 2 .



Εικόνα 2: Οικογένεια RANSAC με αλγορίθμους κατηγοριοποιημένους ανάλογα με τις επιδόσεις και τις βασικές στρατηγικές τους (Xie κ.ά. 2020).

2.1.4 Κατάτμηση βάσει ομαδοποίησης

Οι μέθοδοι που βασίζονται στην ομαδοποίηση χρησιμοποιούνται ευρέως για την κατάτμηση νεφών σημείων χωρίς επίβλεψη. Αυτές οι μέθοδοι δεν βασίζονται σε μια συγκεκριμένη μαθηματική θεωρία. Είναι ένα μείγμα διαφορετικών μεθόδων που μοιράζονται έναν παρόμοιο στόχο, ο οποίος είναι η ομαδοποίηση σημείων με παρόμοια γεωμετρικά χαρακτηριστικά, φασματικά χαρακτηριστικά ή χωρική κατανομή στο ίδιο ομοιογενές μοτίβο. Σε αντίθεση με την ανάπτυξη περιοχών και την προσαρμογή μοντέλων, αυτά τα μοτίβα συνήθως δεν ορίζονται εκ των προτέρων, και έτσι οι αλγόριθμοι που βασίζονται στην ομαδοποίηση μπορούν να χρησιμοποιηθούν για την κατάτμηση ακανόνιστων αντικειμένων, π.χ. βλάστηση. Επιπλέον, τα σημεία εκκίνησης δεν απαιτούνται από τις προσεγγίσεις που βασίζονται στην ομαδοποίηση, σε αντίθεση με τις μεθόδους ανάπτυξης περιοχών. Στο πρώιμο στάδιο, ο K-μέσος (K-means), η μέση μετατόπιση (Mean-shift) και η ασαφής ομαδοποίηση (fuzzy clustering) ήταν οι κύριοι αλγόριθμοι της οικογένειας κατάτμησης νεφών σημείων με βάση την ομαδοποίηση (Xie κ.ά. 2020).

K-μέσος (K-means)

Ο αλγόριθμος K-μέσος είναι ένας βασικός και ευρέως χρησιμοποιούμενος αλγόριθμος ανάλυσης κατά συστάδες. Διαχωρίζει το σύνολο δεδομένων νέφους σημείων σε K μη επισημασμένες κλάσεις. Τα κέντρα ομαδοποίησης των K- μέσων είναι διαφορετικά από τα σημεία εκκίνησης της ανάπτυξης περιοχών. Στον K-μέσο, κάθε σημείο θα πρέπει να συγκρίνεται με κάθε κέντρο συστάδας σε κάθε βήμα επανάληψης και τα

κέντρα συστάδων θα αλλάζουν όταν απορροφάται ένα νέο σημείο. Η διαδικασία του K-μέσου είναι η "ομαδοποίηση" και όχι η "ανάπτυξη". Μερικές από τις εφαρμογές που έχει χρησιμοποιηθεί είναι η κατάτμηση της κόμης ενός δέντρου, η εξαγωγή επίπεδης δομής από στέγες και για την τμηματοποίηση προσόψεων κτιρίων.

Ένα πλεονέκτημα του K-μέσου είναι ότι μπορεί εύκολα να προσαρμοστεί σε όλα τα είδη χαρακτηριστικών και μπορεί να χρησιμοποιηθεί ακόμη και σε πολυδιάστατο χώρο χαρακτηριστικών. Το κύριο μειονέκτημα είναι ότι μερικές φορές είναι δύσκολο να προκαθοριστεί σωστά η τιμή του K (Xie κ.ά. 2020).

Μέση μετατόπιση (Mean-shift)

Σε αντίθεση με τον K-μέσο, ο αλγόριθμος μέση μετατόπιση είναι ένας κλασικός μη παραμετρικός αλγόριθμος ομαδοποίησης και ως εκ τούτου αποφεύγει το πρόβλημα του προκαθορισμένου K στον K-μέσο. Έχει εφαρμοστεί αποτελεσματικά σε δεδομένα που αφορούν αστικό και δασικό έδαφος, όπως επίσης και για εξαγωγή προσόψεων κτιρίων και μεμονωμένων δέντρων. Καθώς τόσο ο αριθμός των συστάδων όσο και το σχήμα κάθε συστάδας είναι άγνωστα, η μετατόπιση του μέσου όρου παρέχει με μεγάλη πιθανότητα υπερτμηματοποιημένο αποτέλεσμα. Ως εκ τούτου, χρησιμοποιείται συνήθως ως βήμα προ-τμηματοποίησης πριν από την κατάτμηση ή τη βελτίωση.

Ασαφής ομαδοποίηση (fuzzy clustering)

Οι αλγόριθμοι ασαφούς ομαδοποίησης είναι βελτιωμένες εκδόσεις του K-μέσου. Ο αλγόριθμος K-μέσος είναι μια απόλυτη μέθοδος συσταδοποίησης, που σημαίνει ότι το βάρος ενός σημείου δείγματος σε ένα κέντρο συστάδας είναι είτε 1 είτε 0. Αντίθετα, οι ασαφείς μέθοδοι χρησιμοποιούν μια όχι τόσο αυστηρή ομαδοποίηση, που σημαίνει ότι ένα σημείο δείγματος μπορεί να ανήκει σε διάφορες ομάδες με ορισμένα μη μηδενικά βάρη.

Στη κατάτμηση νεφών σημείων, δημιουργήθηκε ένα πλαίσιο χωρίς αρχικοποίηση συνδυάζοντας δύο ασαφείς αλγορίθμους, τον αλγόριθμο Fuzzy C-Means (FCM) και τον αλγόριθμο Possibilistic C-Means (PCM). Αυτό το πλαίσιο δοκιμάστηκε σε διάφορα σύνολα δεδομένων όπως εξωτερικοί χώροι με δομές κτιρίων. Τα εν λόγω πειράματα έδειξαν ότι η ασαφής κατάτμηση κατά συστάδες λειτουργούσε σταθερά σε επίπεδες επιφάνειες.

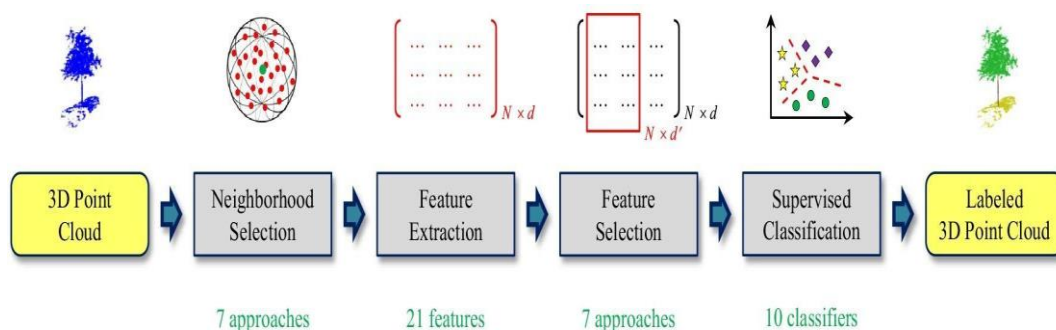
2.2 Τεχνικές Σημασιολογικής Κατάτμησης Νεφών Σημείων

Όπως ήδη έχει αναφερθεί, σημασιολογική κατάτμηση νεφών σημείων θεωρείται η συσχέτιση κάθε σημείου ενός νέφους σημείων με μια σημασιολογική ετικέτα. Παρ' όλο που η διαδικασία της σημασιολογικής κατάτμησης είναι παρόμοια με εκείνη της κατάτμησης νεφών σημείων βάσει της ομαδοποίησης, η ειδοποιός διαφορά τους είναι ότι στην σημασιολογική κατάτμηση δίνεται ερμηνεία σε κάθε σημείο και δεν ομαδοποιούνται μόνο όπως στην απλή κατάτμηση (Xie κ.ά. 2020). Επιπλέον η σημασιολογική κατάτμηση συνήθως πραγματοποιείται μέσω επιβλεπόμενων μεθόδων μηχανικής μάθησης και τα τελευταία χρόνια μέσω της βαθιάς μάθησης.

2.2.1 Κλασσική επιβλεπόμενη μηχανική μάθηση

Πολλοί ερευνητές έχουν αναλύσει διεξοδικά και έχουν συγκρίνει μεταξύ τους τις διάφορες μεθόδους που έχουν αναπτυχθεί για την σημασιολογική κατάτμηση νεφών σημείων οι οποίες αφορούν στην κλασσική επιβλεπόμενη μηχανική μάθηση.

Η επιβλεπόμενη μηχανική μάθηση που εφαρμόζεται στη σημασιολογική κατάτμηση νεφών σημείων μπορεί να χωριστεί σε δύο ομάδες. Η μία ομάδα είναι οι μεμονωμένες σημασιολογικές κατατμήσεις οι οποίες ταξινομούν κάθε σημείο ή κάθε συστάδα σημείων με βάση μόνο τα ατομικά χαρακτηριστικά του, όπως οι «ταξινομητές μέγιστης πιθανοφάνειας» (Maximum Likelihood classifiers) που βασίζονται σε «μοντέλα μίξης Gauss» (GMM), οι «μηχανές διανυσμάτων υποστήριξης» (Support Vector Machines), ένας «καταρράκτης δυαδικών ταξινομητών» (cascade of binary classifiers) και τα «τυχαία δάση» (Random Forest). Η άλλη ομάδα είναι τα στατιστικά μοντέλα πλαισίου, όπως τα «συσχετιστικά και μη συσχετιστικά δίκτυα Μαρκόφ» (Associative and Non-Associative Markov Networks), τα «τυχαία πεδία υπό συνθήκη» (Conditional Random Fields), «απλοποιημένα τυχαία πεδία Μαρκόφ» (Simplified Markov Random Fields), τα οποία αποτελούν διαδικασίες εξαγωγής συμπερασμάτων πολλαπλών σταδίων που εστιάζουν σε στατιστικές και σχεσιακές πληροφορίες για νέφη σημείων σε διαφορετικές κλίμακες και μηχανές χωρικής εξαγωγής συμπερασμάτων που μοντελοποιούν εξαρτήσεις μεσαίας και μεγάλης εμβέλειας που υπάρχουν στα δεδομένα (Xie κ.ά. 2020).



Εικόνα 3: Το πλαίσιο της σημασιολογικής κατάτμησης νεφών σημείων (Weinmann κ.ά. 2015).

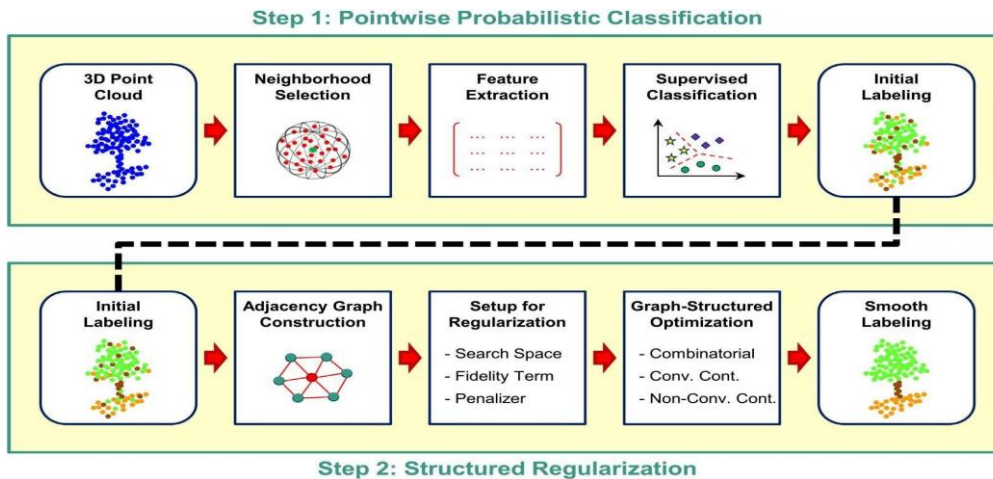
Η γενική διαδικασία της μεμονωμένης κατάτμησης περιλαμβάνει τέσσερα στάδια όπως φαίνεται στην εικόνα 3.

Τα τέσσερα στάδια είναι:

- Επιλογή γειτονιάς
- Εξαγωγή χαρακτηριστικών
- Επιλογή χαρακτηριστικών
- Σημασιολογική κατάτμηση

Για κάθε στάδιο συμπεριλήφθηκαν οι πιο καίριες μέθοδοι και για να συγκριθούν οι αποδόσεις τους δοκιμάστηκαν σε δύο σύνολα δεδομένων. Σύμφωνα με το πείραμα των συγγραφέων (Weinmann κ.ά. 2015), στην μεμονωμένη σημασιολογική κατάτμηση, ο ταξινομητής Random Forest είχε μια καλή αντιστάθμιση μεταξύ ακρίβειας και αποτελεσματικότητας σε δύο σύνολα δεδομένων. Στο πείραμα αναφέρεται ότι χρησιμοποιήθηκε ένας ταξινομητής με τη λεγόμενη "βαθιά μάθηση", αλλά πρόκειται για ένα παλιό νευρωνικό δίκτυο που εμφανίστηκε στην εποχή της απλής μηχανικής μάθησης και όχι για τις πρόσφατες μεθόδους βαθιάς μάθησης που περιγράφονται παρακάτω.

Δεδομένου ότι η μεμονωμένη σημασιολογική κατάτμηση δεν λαμβάνει υπόψη τα χαρακτηριστικά των σημείων που σχετίζονται με το πλαίσιο, οι μεμονωμένοι ταξινομητές λειτουργούν αποτελεσματικά, αλλά δημιουργούν αναπόφευκτο θόρυβο που προκαλεί μη ομαλά αποτελέσματα. Τα στατιστικά μοντέλα πλαισίου μπορούν να μετριάσουν αυτό το πρόβλημα. Τα υπό συνθήκη τυχαία πεδία (Conditional Random Fields) είναι το πιο ευρέως χρησιμοποιούμενο μοντέλο πλαισίου στη σημασιολογική κατάτμηση. Οι Landrieu κ.ά. (2017) πρότειναν ένα νέο πλαίσιο σημασιολογικής κατάτμησης που συνδυάζει την ατομική ταξινόμηση και την ταξινόμηση πλαισίου. Όπως φαίνεται στην εικόνα 4, σε αυτό το πλαίσιο εισήχθη μια στρατηγική βασισμένη σε γράφους για να ξεπεραστεί το πρόβλημα του θορύβου της αρχικής επισήμανσης, από το οποίο η διαδικασία ονομάστηκε δομημένη κανονικοποίηση ή "εξομάλυνση". Συνέκριναν διάφορες μεθόδους μετεπεξεργασίας στις μελέτες τους, οι οποίες απέδειξαν ότι η κανονικοποίηση βελτίωσε πράγματι την ακρίβεια της σημασιολογικής κατάτμησης του νέφους σημείων (Xie κ.ά. 2020).



Εικόνα 4: Το πλαίσιο της σημασιολογικής κατάτμησης νέφους σημείων (Landrieu κ.ά. 2017).

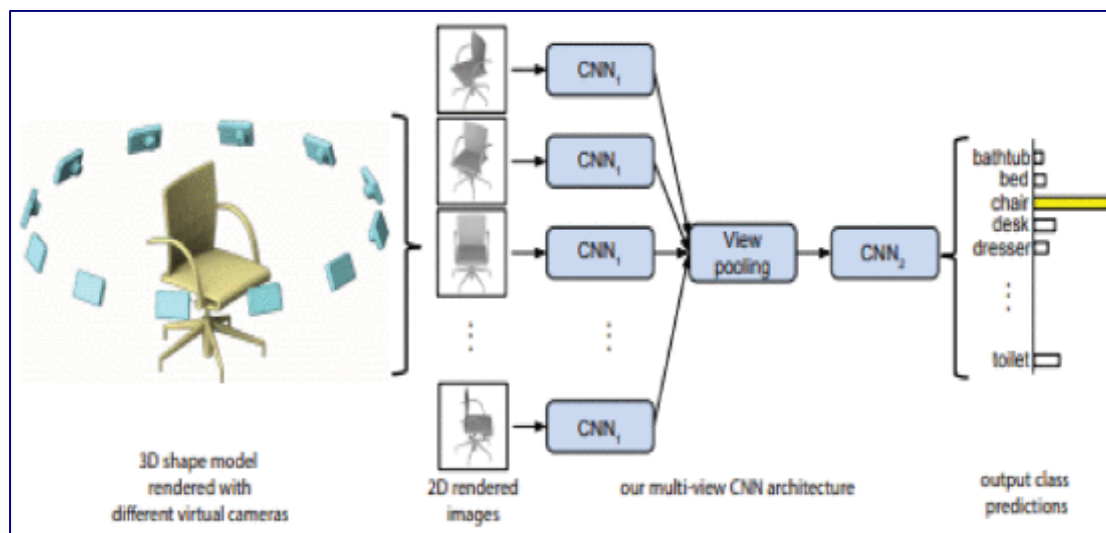
2.2.2 Βαθιά Μηχανική Μάθηση

Η βαθιά μάθηση (deep learning) είναι μια από τις πιο πρόσφατες και γρήγορα εξελισσόμενες τεχνικές στην αναγνώριση προτύπων (pattern recognition), την όραση υπολογιστών (computer vision) και την ανάλυση δεδομένων (data analysis). Τα τελευταία χρόνια, έχει προταθεί μεγάλος αριθμός μοντέλων που χρησιμοποιούν αυτές τις μεθόδους για την επεξεργασία του νέφους σημείων. Πριν εφαρμοστεί σε τρισδιάστατα δεδομένα, η βαθιά μάθηση εμφανίστηκε ως αποτελεσματική δύναμη σε μια ποικιλία εργασιών στη δισδιάστατη όραση υπολογιστών και επεξεργασία εικόνας, όπως η αναγνώριση εικόνων, η ανίχνευση αντικειμένων και η σημασιολογική κατάτμηση. Από το 2015 έχει προσελκύσει μεγαλύτερο ενδιαφέρον στην τρισδιάστατη ανάλυση, με αφορμή την ιδέα της πολλαπλής προβολής που προτάθηκε από τους Su κ.ά. (Su κ.ά. 2015) και το 3D Νευρωνικό Συνελκτικό Δίκτυο (Convolutional Neural Network) με βάση τα ογκομετρικά στοιχεία (voxel) από τους (Maturana και Scherer 2015).

Οι συνήθεις συνελίξεις που σχεδιάστηκαν αρχικά για κανονικοποιημένες εικόνες (raster) δεν μπορούν εύκολα να εφαρμοστούν άμεσα στη σημασιολογική κατάτμηση ενός νέφους σημείων, καθώς το νέφος σημείων είναι αταξινόμητο και μη δομημένο κανονικοποιημένο. Έτσι, για να λυθεί αυτό το πρόβλημα, καθίσταται απαραίτητος ο μετασχηματισμός του ακατέργαστου νέφους σημείων. Ανάλογα με τη μορφή των δεδομένων που εισάγονται στα νευρωνικά δίκτυα, οι προσεγγίσεις της σημασιολογικής κατάτμησης που βασίζονται στη βαθιά μάθηση μπορούν να χωριστούν σε τρεις κατηγορίες: βασισμένες σε πολλαπλές εικόνες, βασισμένες σε ογκομετρικά εικονοστοιχεία (voxel) και βασισμένες σε σημεία.

2.2.2.1 Σημασιολογική κατάτμηση βασισμένη στη μέθοδο πολλαπλών προβολών (Multiview-based)

Μια από τις πρώτες λύσεις για την εφαρμογή της βαθιάς μάθησης είναι η μείωση της διαστατικότητας. Αυτό σημαίνει ότι τα τρισδιάστατα δεδομένα αναπαρίστανται από εικόνες 2D πολλαπλών προβολών, οι οποίες μπορούν να υποβληθούν σε επεξεργασία με βάση τα 2D συνελκτικά νευρωνικά δίκτυα (convolutional neural network). Στη συνέχεια, τα αποτελέσματα της ταξινόμησης μπορούν να αποκατασταθούν σε 3D. Σ' αυτό βασίστηκαν οι Su κ.ά (2015) και πρότειναν ένα συνελκτικό νευρωνικό δίκτυο πολλαπλών προβολών βασισμένο σε εικόνες, το οποίο προωθεί την ανάπτυξη επεξεργασίας τρισδιάστατων δεδομένων. Η κύρια ιδέα αυτής της μεθόδου είναι να προβάλλεται το τρισδιάστατο νέφος σημείων σε ορισμένες δισδιάστατες εικόνες από πολλαπλές προβολές και να χρησιμοποιείται το συνελκτικό νευρωνικό δίκτυο για την εξαγωγή των χαρακτηριστικών για κάθε προβολή χρησιμοποιώντας τις μεθόδους επεξεργασίας εικόνας και έπειτα να συγκεντρώνονται όλα τα χαρακτηριστικά που εξάγονται από διαφορετικές προβολές μέσω του επιπέδου συγκέντρωσης προβολών. Τέλος, τα συγκεντρωτικά χαρακτηριστικά εισάγονται στο συνελκτικό νευρωνικό δίκτυο για επεξεργασία, λαμβάνοντας έτσι τα αποτελέσματα της ταξινόμησης και της κατάτμησης Su κ.ά (Su κ.ά 2015)-(Zhang κ.ά. 2019). Η εικόνα 5 αποτελεί πλήρη απεικόνιση αυτής της μεθόδου.

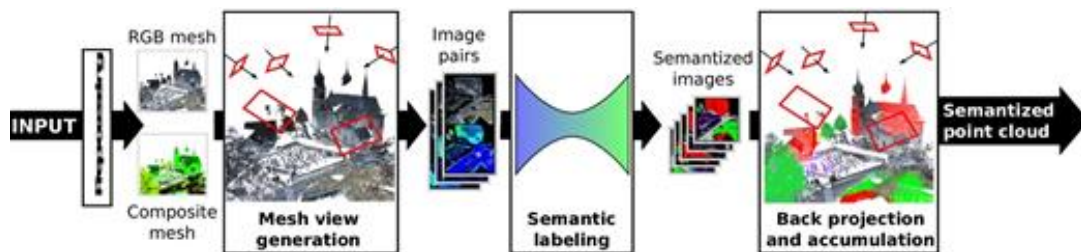


Εικόνα 5: Το πλαίσιο του νευρωνικού δικτύου πολλαπλών όψεων (MVCNN (Su κ.ά 2015))

Οι μέθοδοι που βασίζονται σε πολλαπλές προβολές παρόλο που έχουν επιλύσει καλά τα προβλήματα δόμησης των δεδομένων νέφους σημείων, έχουν ακόμα δύο σοβαρές ελλείψεις. Η μία είναι ότι προκαλούν πολυάριθμους περιορισμούς και απώλεια των γεωμετρικών δομών καθώς οι δισδιάστατες εικόνες πολλαπλών προβολών είναι απλά μια προσέγγιση των τρισδιάστατων σκηνών. Αυτό έχει ως αποτέλεσμα σύνθετες εργασίες όπως είναι η σηματολογική κατάτμηση νεφών σημείων να μην έχουν τα επιθυμητά αποτελέσματα και οι επιδόσεις τους να είναι περιορισμένες και μη ικανοποιητικές. Η δεύτερη έλλειψη των μεθόδων αυτών είναι ότι επειδή οι εικόνες πολλαπλών όψεων πρέπει να καλύπτουν όλους τους χώρους που περιέχουν σημεία, είναι δύσκολο για μεγάλες και πολύπλοκες σκηνές να επιλεγθούν αρκετά κατάλληλα σημεία θέασης για προβολή. Γι' αυτό το λόγο λίγες μελέτες βρέθηκαν να

χρησιμοποιούν αυτή τη μέθοδο βαθιάς μηχανικής μάθησης βασισμένη σε πολλαπλές προβολές (Xie κ.ά. 2020).

Μια μελέτη που βασίστηκε σε αυτή τη μέθοδο είναι των Boulch κ.ά. (2018). Αυτοί ανέπτυξαν ένα πλαίσιο το οποίο εφαρμόζει τα συνελκτικά νευρωνικά δίκτυα σε πολλαπλές προβολές δισδιάστατων εικόνων (ή στιγμιότυπων) του νέφους σημείων. Το SnapNet αποτελείται από 3 βασικά βήματα. Το πρώτο βήμα είναι η επιλογή κατάλληλων στιγμιότυπων του νέφους σημείων και η δημιουργία δύο τύπων εικόνων, μια προβολή RGB και μια σύνθετη προβολή βάθους που περιέχει τα γεωμετρικά χαρακτηριστικά. Στη συνέχεια, εκτελείται μια επισήμανση κατά εικονοστοιχείο κάθε ζεύγους δισδιάστατων στιγμιότυπων χρησιμοποιώντας πλήρως συνελκτικά δίκτυα. Τέλος, εκτελείται μια γρήγορη οπισθοπροβολή των προβλέψεων ετικέτας στο τρισδιάστατο χώρο χρησιμοποιώντας αποδοτική αποθήκευση για την ετικέτα κάθε τρισδιάστατου σημείου. Με αυτόν τον τρόπο επιτυγχάνεται η προβολή των αποτελεσμάτων της δισδιάστατης σημασιολογικής κατάτμησης στον τρισδιάστατο χώρο, ώστε να αποκτηθεί η σημασιολογική κατάτμηση του νέφους σημείων. Στην εικόνα 6 παρουσιάζεται η ροή εργασιών του SnapNet.



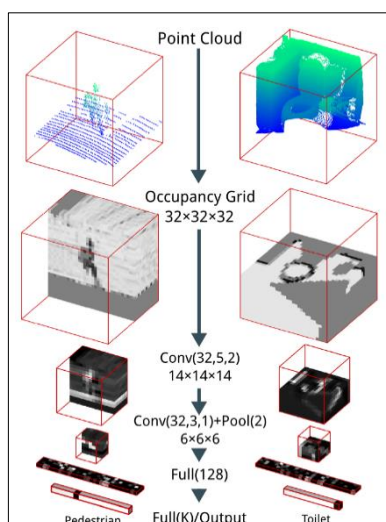
Εικόνα 6: Η ροή εργασιών της μεθόδου SnapNet (Boulch κ.ά 2018).

2.2.2.2 Σημασιολογική κατάτμηση βασισμένη σε ογκομετρικά εικονοστοιχεία (voxel):

Ο συνδυασμός ογκομετρικών στοιχείων με τα τρισδιάστατα συνελκτικά νευρωνικά δίκτυα είναι άλλη μία πρώιμη προσέγγιση στη σημασιολογική κατάτμηση των νεφών σημείων που βασίζεται στη βαθιά μάθηση. Η ογκομετροποίηση (voxelization) επιλύει τόσο τα μη ταξινομημένα όσο και τα μη δομημένα προβλήματα του ακατέργαστου νέφους σημείων. Με την ογκομετροποίηση μετασχηματίζονται τα μη δομημένα νέφη σημείων και μετατρέπονται σε συμμετρικά ογκομετρικά πλέγματα πληρότητας (regular volumetric occupancy grid). Στη συνέχεια, με την χρήση νευρωνικών δικτύων επιτυγχάνεται η εκμάθηση των χαρακτηριστικών τους και ως εκ τούτου η σημασιολογική κατάτμηση των νεφών σημείων (Xie κ.ά. 2020).

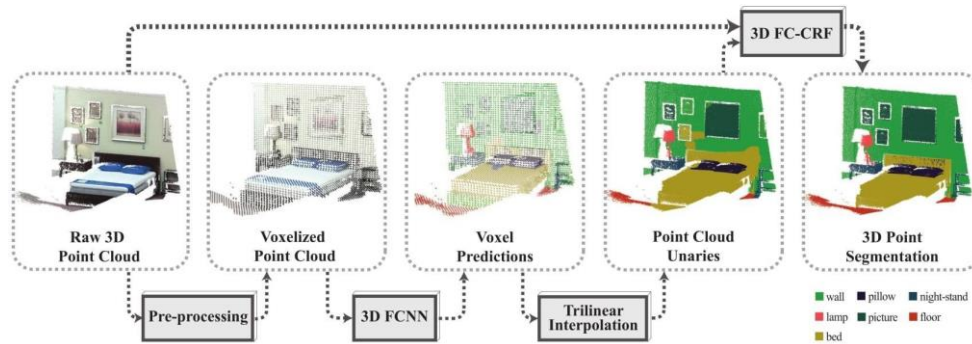
Το πιο γνωστό τρισδιάστατο συνελκτικό νευρωνικό δίκτυο με βάση τα ογκομετρικά εικονοστοιχεία είναι το VoxNet (Maturana και Scherer 2015). Σύμφωνα με αυτό, τα μη δομημένα γεωμετρικά δεδομένα μετατρέπονται σε ένα συμμετρικό τρισδιάστατο πλέγμα πάνω στο οποίο μπορούν να εφαρμοστούν οι τυπικές λειτουργίες ενός συνελκτικού νευρωνικού δικτύου. Στη συνέχεια χρησιμοποιείται ένα τρισδιάστατο συνελκτικό νευρωνικό δίκτυο με το οποίο πραγματοποιείται η πρόβλεψη της ετικέτας απευθείας από το πλέγμα, δηλαδή η πρόβλεψη για το τι είναι η κάθε κλάση. Η μέθοδος αυτή επιλύει το γεγονός ότι τα νέφη σημείων είναι μη δομημένα γεωμετρικά δεδομένα, αλλά συνεχίζει και έχει αρκετές αδυναμίες. Αυτές είναι η χαμηλή αποτελεσματικότητα

της διάταξης του πλέγματος voxel που προκαλείται από τη σπανιότητα των σημειακών νεφών, η μεγάλη μνήμη που καταλαμβάνεται κατά τη διαδικασία υπολογισμού, ο μεγάλος χρόνος εκπαίδευσης και το πρόβλημα της απώλειας πληροφορίας.



Εικόνα 7: Η αρχιτεκτονική του Voxelnet (Maturana και Scherer 2015).

Ένα πιο εξελιγμένο πλαίσιο συνελκτικού νευρωνικού δικτύου είναι το SegCloud (Tchapmi κ.ά 2017). Είναι ένα ολοκληρωμένο πλαίσιο που συνδυάζει τα πλεονεκτήματα των τρισδιάστατων πλήρως συνελκτικών νευρωνικών δικτύων (3D Fully Convolutional Neural Networks), της τριγραμμικής παρεμβολής (Trilinear Interpolation) και πλήρως συνδεδεμένα τυχαία πεδία υπό συνθήκη (Fully Connected Conditional Random Fields) για να φέρει εις πέρας το έργο της σημασιολογικής κατάτμησης. Στην εικόνα 8 παρουσιάζεται το πλαίσιο του SegCloud, το οποίο παρέχει επίσης ένα βασικό εργαλείο σημασιολογικής κατάτμησης με βάση τα voxel. Στο SegCloud, το βήμα προεπεξεργασίας είναι η ογκομετροποίηση των ακατέργαστων νεφών σημείων. Στη συνέχεια εφαρμόζεται ένα τρισδιάστατο πλήρως συνελκτικό νευρωνικό δίκτυο για τη δημιουργία ετικετών ογκομετρικών στοιχείων (voxel) με μειωμένη δειγματοληψία. Στη συνέχεια, ένα επίπεδο τριγραμμικής παρεμβολής χρησιμοποιείται για τη μεταφορά των ετικετών ογκομετρικών στοιχείων (voxel) σε ετικέτες τρισδιάστατων σημείων. Τέλος, χρησιμοποιείται μια τρισδιάστατη πλήρως συνδεδεμένη μέθοδος τυχαίων πεδίων υπό συνθήκη για την κανονικοποίηση προηγούμενων αποτελεσμάτων σημασιολογικής κατάτμησης και την απόκτηση τελικών αποτελεσμάτων. Το SegCloud δεν έλαβε κανένα μέτρο για τη βελτιστοποίηση του προβλήματος της υψηλής υπολογιστικής απόδοσης και μνήμης από τα voxels σταθερού μεγέθους γι αυτό και με την εμφάνιση πιο προηγμένων μεθόδων, το SegCloud έπεσε σε δυσμένεια τα τελευταία χρόνια.



Εικόνα 8: Το πλαίσιο SegCloud (Tchapmi κ.ά 2017)

2.2.2.3 Άμεση επεξεργασία δεδομένων νέφους σημείων (Directly process point cloud data)

Όπως διαπιστώθηκε στην παραπάνω ανάλυση, υπάρχουν σοβαροί περιορισμοί τόσο στις μεθόδους που βασίζονται σε πολλαπλές προβολές (multiview-based methods) όσο και στις μεθόδους που βασίζονται σε ογκομετρικά στοιχεία (voxel-based methods). Αυτό έγκειται στο γεγονός ότι τα νέφη σημείων έχουν μια ακανόνιστη δομή και οι παραπάνω μέθοδοι αδυνατούν να την αναλύσουν. Γι' αυτό το λόγο αναπτύχθηκαν κάποιες μέθοδοι βάσει των οποίων η σημασιολογική κατάτμηση των νεφών σημείων γίνεται απευθείας στα σημεία. Σε αντίθεση με τη χρήση ξεχωριστής λειτουργίας προμετασχηματισμού στις περιπτώσεις που βασίζονται σε πολλαπλές προβολές και σε ογκομετρικά στοιχεία, σε αυτές τις προσεγγίσεις η κανονικοποίηση είναι δεσμευτική με την αρχιτεκτονική του νευρωνικού δικτύου.

Το PointNet (Qi κ.ά. 2016) είναι ένα πρωτοποριακό πλαίσιο βαθιάς μάθησης το οποίο έχει εκτελεστεί απευθείας σε σημεία. Διαφορετικά με τα πρόσφατα δημοσιευμένα δίκτυα νέφους σημείων, δεν υπάρχει τελεστής συνέλιξης στο PointNet. Η βασική αρχή του PointNet είναι:

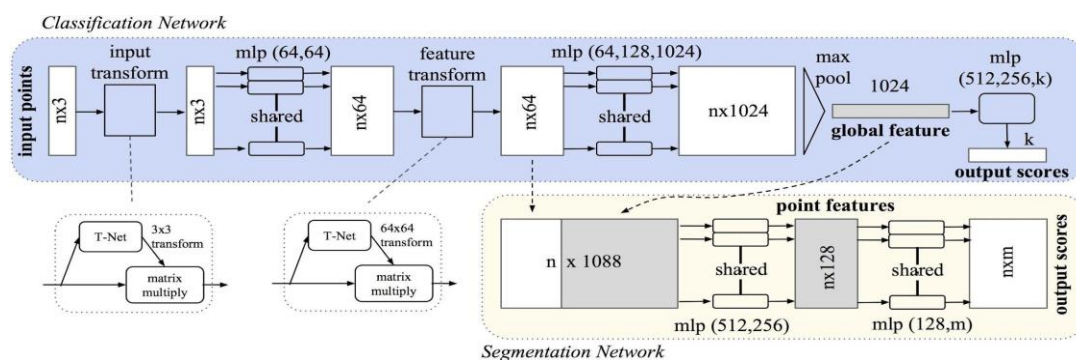
$$f(\{x_1, \dots, x_n\}) \approx g(h(x_1), \dots, h(x_n))$$

where $f : 2^{\mathbb{R}^N} \rightarrow \mathbb{R}$ and $h : \mathbb{R}^N \rightarrow \mathbb{R}^K$. $g :$

$\underbrace{\mathbb{R}^K \times \dots \times \mathbb{R}^K}_n \rightarrow \mathbb{R}$ είναι μια συμμετρική διάταξη που επιλύει το πρόβλημα της διάταξης των νεφών σημείων.

Όπως φαίνεται και στην εικόνα 9, το Pointnet χρησιμοποιεί πολυεπίπεδα νευρωνικά δίκτυα (Multilayer Perceptrons) για να προσεγγίσει την παράμετρο h , η οποία αντιπροσωπεύει τα τοπικά χαρακτηριστικά ανά σημείο δηλαδή τα χαρακτηριστικά που αντιστοιχούν σε κάθε σημείο. Τα παγκόσμια χαρακτηριστικά g των σημείων του συνόλου των δεδομένων είναι το άθροισμα όλων των τοπικών χαρακτηριστικών ανά σημείο μέσω μια συμμετρικής συνάρτησης, της μέγιστης συγκέντρωσης (max pooling). Για την εργασία της ταξινόμησης, οι βαθμολογίες εξόδου για k κλάσεις μπορούν να παραχθούν από την λειτουργία των πολυεπίπεδων νευρωνικών δικτύων (multilayer perceptrons) στα παγκόσμια χαρακτηριστικά. Για την εργασία της σημασιολογικής κατάτμησης νεφών σημείων, εκτός από τα παγκόσμια χαρακτηριστικά

απαιτούνται και τα τοπικά χαρακτηριστικά ανά σημείο. Το PointNet συνδυάζει συγκεντρωτικά παγκόσμια χαρακτηριστικά και τοπικά χαρακτηριστικά ανά σημείο σε συνδυασμένα χαρακτηριστικά σημείου. Στη συνέχεια, νέα χαρακτηριστικά ανά σημείο εξάγονται από τα συνδυασμένα χαρακτηριστικά σημείων μέσω πολυεπίπεδων νευρωνικών δικτύων. Βάσει αυτών, προβλέπονται οι σημασιολογικές ετικέτες.



Εικόνα 9: Το πλαίσιο του PointNet (Qi κ.ά. 2016).

Παρόλο που όλο και περισσότερα πρόσφατα δημοσιευμένα δίκτυα υπερτερούν του PointNet σε διάφορα σύνολα δεδομένων αναφοράς, το PointNet εξακολουθεί να αποτελεί τη βασική επιλογή για την έρευνα σημασιολογικής κατάτμησης νεφών σημείων. Το αρχικό PointNet δεν χρησιμοποιεί πληροφορίες τοπικής δομής σε γειτονικά σημεία. Σε μια άλλη μελέτη, οι Qi κ. ά. (2016) χρησιμοποίησαν ένα ιεραρχικό νευρωνικό δίκτυο για να συλλάβουν τα τοπικά γεωμετρικά χαρακτηριστικά ώστε να βελτιώσουν το βασικό μοντέλο PointNet και πρότειναν το PointNet++. Αντλώντας έμπνευση από το PointNet/PointNet++, οι μελέτες σχετικά με την τρισδιάστατη βαθιά μάθηση επικεντρώνονται στην αύξηση των χαρακτηριστικών, ιδίως στα τοπικά χαρακτηριστικά και τις σχέσεις μεταξύ των σημείων, αξιοποιώντας τη γνώση από άλλα πεδία για τη βελτίωση της απόδοσης των βασικών αλγορίθμων PointNet/PointNet++. Για παράδειγμα, οι Engelmann κ.ά. το 2017 χρησιμοποίησαν δύο επεκτάσεις στο PointNet για να ενσωματώσει χωρικό πλαίσιο μεγαλύτερης κλίμακας. Οι Wang κ.ά. το 2018 θεώρησαν ότι η έλλειψη τοπικών χαρακτηριστικών εξακολουθούσε να αποτελεί πρόβλημα στο PointNet++, δεδομένου ότι παραμελούσε τις γεωμετρικές σχέσεις μεταξύ ενός σημείου και των γειτονικών του. Για να ξεπεραστεί αυτό το πρόβλημα, πρότειναν το Dynamic Graph Convolutional Neural Networks. Σε αυτό το δίκτυο, οι συγγραφείς σχεδίασαν μια διαδικασία που ονομάζεται EdgeConv για να εξάγουν χαρακτηριστικά ακμών διατηρώντας παράλληλα το αναλλοίωτο της μετατροπής. Εμπνευσμένοι από την ιδέα του μηχανισμού της προσοχής, οι Wang κ. ά. σχεδίασαν μια συνέλιξη Graph Attention Convolution (GAC), της οποίας οι πυρήνες μπορούν να προσαρμόζονται δυναμικά στη δομή ενός αντικειμένου. Το GAC μπορεί να συλλάβει τα δομικά χαρακτηριστικά των νεφών σημείων, αποφεύγοντας παράλληλα τη “μόλυνση” των χαρακτηριστικών μεταξύ των αντικειμένων. Για να αξιοποιήσουν περισσότερα χαρακτηριστικά ακμών, οι Landrieu και Simonovsky (2018) εισήγαγαν το SuperPoint Graph (SPG), προσφέροντας συμπαγή και πλούσια αναπαράσταση των σχέσεων πλαισίου μεταξύ των τμημάτων του αντικειμένου και όχι των σημείων. Η κατάτμηση του υπερσημείου, δηλαδή ενός μικρού τμήματος του νέφους σημείων που

αντιστοιχεί σε ένα απλό γεωμετρικό σχήμα μπορεί να θεωρηθεί ως ένα μη σημασιολογικό βήμα προ-τμηματοποίησης. Μετά την κατασκευή SPG, κάθε υπερσημείο ενσωματώνεται σε ένα βασικό δίκτυο PointNet και στη συνέχεια βελτιώνεται σε Gated Recurrent Units (GRUs) για σημασιολογική κατάτμηση του νέφους σημείων. Επωφελούμενη από την πλούσια σε πληροφορίες υποδειγματοληψία, η SPG είναι ιδιαίτερα αποτελεσματική για σύνολα δεδομένων μεγάλου όγκου. Επίσης, προκειμένου να ξεπεραστεί το μειονέκτημα της απουσίας τοπικών χαρακτηριστικών που αντιπροσωπεύονται από γειτονικά σημεία στο PointNet, το 3P-Recurrent Neural Network υιοθέτησε μια μονάδα Pointwise Pyramid Pooling (3P) για να συλλάβει τα τοπικά χαρακτηριστικά κάθε σημείου. Επιπλέον, χρησιμοποίησε ένα μοντέλο επαναλαμβανόμενου νευρωνικού δικτύου (RNN) δύο κατευθύνσεων για την ενσωμάτωση του πλαισίου μεγάλης εμβέλειας σε εργασίες σημασιολογικής κατάτμησης νεφών σημείων. Η τεχνική 3P-RNN αύξησε τη συνολική ακρίβεια με αμελητέα επιπλέον επιβάρυνση. Οι Komarichev κ.ά. το 2019 εισήγαγαν μια δακτυλιοειδή συνέλιξη, η οποία θα μπορούσε να συλλάβει την τοπική γειτονιά καθορίζοντας τις δομές και τις κατευθύνσεις σε σχήμα δακτυλίου στον υπολογισμό και να προσαρμοστεί στη γεωμετρική μεταβλητότητα και την επεκτασιμότητα σε επίπεδο επεξεργασίας σήματος. Λόγω του γεγονότος ότι η αναζήτηση K-κοντινότερου γείτονα στο PointNet++ μπορεί να οδηγήσει στο να εμπίπτουν οι K γείτονες σε έναν προσανατολισμό, οι Jiang κ.ά. το 2018 σχεδίασαν το PointSIFT για να καταγράφουν τοπικά χαρακτηριστικά από οκτώ προσανατολισμούς. Σε ολόκληρη την αρχιτεκτονική, η μονάδα PointSIFT επιτυγχάνει πολυδιάστατη αναπαράσταση με τη στοίβαξη πολλών μονάδων κωδικοποίησης προσανατολισμού (Orientation-Encoding). Η ενότητα PointSIFT μπορεί να ενσωματωθεί σε όλα τα είδη αρχιτεκτονικών τρισδιάστατης βαθιάς μάθησης με βάση το PointNet για να βελτιώσει την ικανότητα αναπαράστασης των τρισδιάστατων σχημάτων. Βασισμένο στο PointNet++, το PointWeb χρησιμοποίησε την ενότητα Adaptive Feature Adjustment (AFA) για να βρει την αλληλεπίδραση μεταξύ των σημείων. Ο στόχος της AFA είναι επίσης να συλλάβει και να συγκεντρώσει τα τοπικά χαρακτηριστικά των σημείων (Xie κ.ά. 2020).

Όπως φαίνεται ένας αυξανόμενος αριθμός ερευνητών έχει επιλέξει μια εναλλακτική λύση του PointNet χρησιμοποιώντας τη συνέλιξη ως θεμελιώδες και σημαντικό συστατικό. Η έρευνα σχετικά με τη σημασιολογική κατάτμηση νεφών σημείων που βασίζεται στη βαθιά μάθηση βρίσκεται ακόμη σε εξέλιξη. Νέες ιδέες και προσεγγίσεις συνεχίζουν να εμφανίζονται σχετικά με το θέμα που βασίζονται στη βαθιά μάθηση. Παρ' όλα αυτά τα μέχρι τώρα επιτεύγματα έχουν αποδείξει ότι υπάρχει μια μεγάλη πρόοδος για την ακρίβεια της τρισδιάστατης σημασιολογικής κατάτμησης νεφών σημείων.

3. ΕΦΑΡΜΟΓΕΣ ΣΗΜΑΣΙΟΛΟΓΙΚΗΣ ΚΑΤΑΤΜΗΣΗΣ

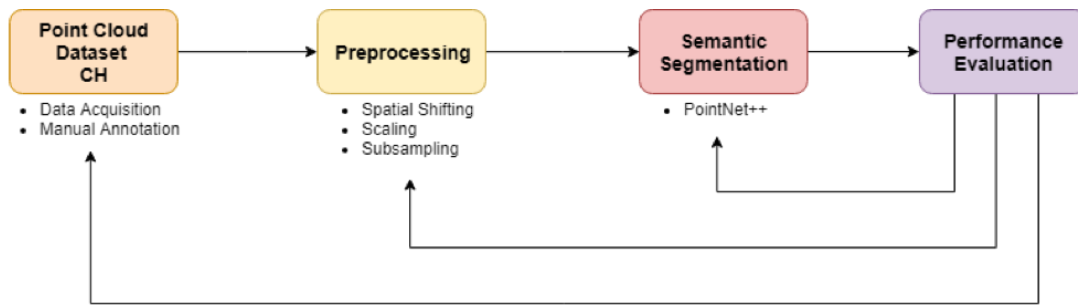
Στο προηγούμενο κεφάλαιο διαχωρίστηκαν οι τρόποι σημασιολογικής κατάτμησης ενός νέφους σημείου με βάση την μεθοδολογία. Ωστόσο μπορούν να διαχωριστούν και ως προς τα δεδομένα, δηλαδή ως προς το τι απεικονίζει το νέφος σημείων. Έχει παρατηρηθεί ότι οι μελέτες για την σημασιολογική κατάτμηση εφαρμόζονται κατά κύριο λόγο σε δυο κατηγορίες δεδομένων, τα δεδομένα πολιτισμικής κληρονομιάς και τα δεδομένα αστικού τύπου. Στο παρόν κεφάλαιο παρουσιάζεται μια μέθοδος σημασιολογικής κατάτμησης για κάθε κατηγορία.

3.1 Πολιτισμική Κληρονομιά.

Η πολιτισμική κληρονομιά αποτελεί ένα σημαντικό κομμάτι της ανθρώπινης ύπαρξης καθώς αναφέρεται στο παρελθόν του ανθρώπου και στα κατορθώματά του. Αυτό σημαίνει ότι και τα αντικείμενα που αφορούν στην πολιτισμική κληρονομιά παρουσιάζουν μεγάλη ποικιλία ως προς τη φύση, το μέγεθος και την πολυπλοκότητά τους. Για τον λόγο αυτό, η τρισδιάστατη σάρωση και η μοντελοποίηση χώρων και αντικειμένων πολιτισμικής κληρονομιάς έχουν αυξηθεί σημαντικά τα τελευταία χρόνια. Πλέον έχει αυξηθεί και η ανάγκη να αναλυθούν αυτά τα αντικείμενα ώστε να μελετηθούν καλύτερα. Σε αυτό συμβάλλει η σημασιολογική κατάτμηση μέσω της οποίας επιτρέπεται ο κατακερματισμός πολύπλοκων αρχιτεκτονικών σε μεμονωμένα στοιχεία. Πολλοί επιστήμονες έχουν υλοποιήσει διάφορες μεθόδους για να το πετύχουν αυτό. Στη παρούσα εργασία παρουσιάζεται η μέθοδος των Malinverni κ.ά [29] η οποία είναι μια μέθοδος για την αυτόματη επισήμανση και ομαδοποίηση ενός νέφους σημείων με βάση μια προσέγγιση βαθιάς μάθησης με επίβλεψη, χρησιμοποιώντας ένα νευρωνικό δίκτυο τελευταίας τεχνολογίας που ονομάζεται PointNet++.

3.1.1 Μεθοδολογία

Παρά το γεγονός ότι είναι γνωστές και άλλες μέθοδοι, οι Malinverni κ.ά (2019) επέλεξαν την PointNet++, καθώς έχει επιτύχει σημαντικά αποτελέσματα για την ταξινόμηση και την τμηματοποίηση τρισδιάστατων νεφών σημείων. Στην εργασία τους το PointNet++ έχει δοκιμαστεί και βελτιωθεί, εκπαιδεύοντας το δίκτυο με σχολιασμένα νέφη σημείων, δηλαδή ορίζοντας τα αντικείμενα που υπάρχουν στην σκηνή, τα οποία προέρχονται από μια πραγματική έρευνα και αξιολογώντας τον τρόπο με τον οποίο αλλάζει η απόδοση ανάλογα με τα δεδομένα εκπαίδευσης εισόδου. Μπορεί να έχει μεγάλο ενδιαφέρον για την ερευνητική κοινότητα που ασχολείται με τη σημασιολογική κατάτμηση νεφών σημείων, καθώς δημοσιοποιεί ένα σύνολο δεδομένων με ετικέτες των στοιχείων πολιτισμικής κληρονομιάς για περαιτέρω δοκιμές (Malinverni κ.ά. 2019).



Εικόνα 10: Ροή εργασιών μελέτης των Malinverni κ.ά (2019).

Η ροή εργασιών που υλοποιήθηκε στη μελέτη αποτελείται από 4 κύρια βήματα όπως φαίνονται και στην εικόνα 10. Αυτά είναι η δημιουργία του συνόλου των δεδομένων, η προεπεξεργασία, η σημασιολογική κατάτμηση με PointNet++ και τέλος η αξιολόγηση αποτελεσμάτων.

3.1.1.1 Δημιουργία συνόλου δεδομένων

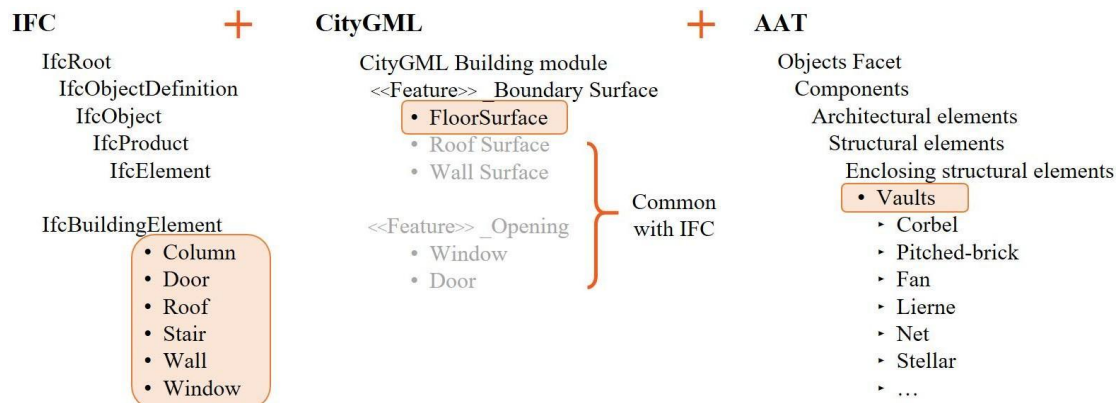
Σύμφωνα με την μελέτη (Malinverni κ.ά. 2019) κατά τη φάση της λήψης, τα νέφη σημείων μπορούν να ληφθούν από διαφορετικά όργανα και κάθε όργανο θα παράγει νέφη με μεταβλητή πυκνότητα σημείων. Γι' αυτό επιλέχθηκαν 4 διαφορετικά μνημεία από τα οποία δημιουργήθηκαν 4 διαφορετικά νέφη σημείων ώστε να υπάρχουν τόσο εσωτερικές όσο και εξωτερικές σκηνές και να αναπτυχθεί μια μέθοδος κατάλληλη για τους περισσότερους ιστορικούς αρχιτεκτονικούς χώρους. Η συλλογή δεδομένων έγινε τόσο με TLS (FARO CAM2) όσο και με UAV (Phantom 4 Pro) ώστε να ενσωματωθούν δύο διαφορετικά νέφη σημείων, το επίγειο κυρίως για τους τοίχους και το εναέριο για τις στέγες.

Τα νέφη σημείων που δημιουργήθηκαν αποτελούνται από εκατομμύρια σημεία, ένας μεγάλος αριθμός σημείων που πρέπει να υποβληθούν σε επεξεργασία στις επόμενες φάσεις. Επιπλέον, με βάση τα όργανα που χρησιμοποιήθηκαν, υπάρχει η δυνατότητα να έχουμε διαφορετικά χαρακτηριστικά για κάθε σημείο. Εκτός από τις γεωγραφικές συντεταγμένες, θα μπορούσαν να έχουν:

- Χρώμα RGB,
- XYZ ,
- Ανακλαστικότητα (“απάντηση” που δίνει το υλικό όταν το χτυπάει το λέιζερ).

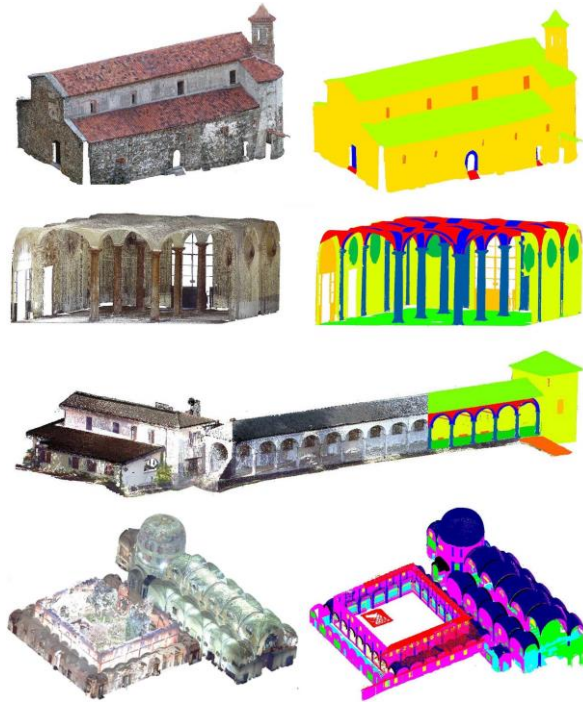
Στη διαδικασία επισήμανσης, η πρώτη φάση αφορά στον προσδιορισμό κάθε αντικείμενου της επίκτητης σκηνής. Επί του παρόντος, αυτή η φάση γίνεται από έναν ανθρώπινο χειριστή, χρησιμοποιώντας διάφορα εργαλεία λογισμικού, όπως, για παράδειγμα, το Autodesk Recap 360. Κάθε σκηνή είναι κατακερματισμένη σε διάφορα αντικείμενα και κάθε αντικείμενο διαθέτει μια τάξη. Προκειμένου να αναπτυχθεί μια ενοποιημένη μέθοδος για την ταξινόμηση αρχιτεκτονικών στοιχείων, έχουν εξεταστεί ορισμένα υπάρχοντα πρότυπα. Συγκεκριμένα, τα πρότυπα .IFC και CityGML καταγράφηκαν. Το πρώτο προορίζεται για το πεδίο AEC (αρχιτεκτονική κατασκευή) με αρχιτεκτονική κλίμακα, ενώ το δεύτερο είναι κυρίως για αστικά και εδαφικά δεδομένα, αλλά, προβλέπει επίσης συγκεκριμένα μέρη για τα στοιχεία κτιρίων. Στο αρχικό στάδιο αυτής της έρευνας που παρουσιάζεται έχει καθιερωθεί για να

ταξινομήσει τα νέφη σημείων σύμφωνα με ένα CityGML LOD 3, επομένως δεν είναι ακόμη ένα μέρος της κατάτμησης, αλλά ακόμα διαφορετικά αρχιτεκτονικά στοιχεία. Επιπλέον, καθώς οι τάξεις που συμπεριλήφθηκαν σε αυτά τα δύο πρότυπα δεν ήταν αρκετές για να περιγράψουν σωστά μια πολιτισμική κληρονομιά, εξετάστηκε το AAT (Art & Architecture Thesaurus) από το Ινστιτούτο Getty. Με αυτό το σκοπό, έχουν επιλεγεί 9 κατηγορίες (Εικόνα 11), συν ένα άλλο που ορίζεται ως "άλλο" με όλα τα μη ταξινομημένα σημεία, που δεν ανήκουν στα προηγούμενα (π.χ. πίνακες, βωμοί, πάγκοι, αγάλματα ...).



Εικόνα 11: Γενικό σχήμα της ενοποιημένης μεθόδου για τον ορισμό και την ταξινόμηση του αρχιτεκτονικού στοιχείου (Malinverni κ.ά. 2019).

Συνεπώς σύμφωνα με την μελέτη, είναι δυνατό να δημιουργηθεί ένας φάκελος για κάθε τάξη και να αποθηκευτούν αντικείμενα της ίδιας κλάσης. Ακόμη και το όνομα πρέπει να μορφοποιηθεί με τέτοιο τρόπο ώστε να αναγνωρίζει τον τύπο του αντικειμένου, την κλάση και τις υποκατηγορίες που θα μπορούσαν να χρησιμοποιηθούν στο μέλλον (παράδειγμα: φάκελος door, αρχείο door-entrancedoor.txt). Σε κάθε αρχείο, όλες οι πληροφορίες του αντικειμένου (π.χ. συντεταγμένες, χρώμα κ.ά) αποθηκεύονται με τέτοιο τρόπο ώστε η επιλογή του οποίου η χρήση θα συμβεί μόνο κατά τη διάρκεια της κατάτμησης. Μια πιο καθαρή απεικόνιση της διαδικασίας που υιοθετήθηκε για την επισήμανση δεδομένων φαίνεται στην εικόνα 12.



Εικόνα 12: Οπτικοποίηση του ορισμού των αρχιτεκτονικών τάξεων. Στα αριστερά το σαρωμένο αντικείμενο, στα δεξιά κάθε συγκεκριμένη κλάση.

3.1.1.2 Προεπεξεργασία

Σε αυτή τη φάση, το αρχικό σύνολο δεδομένων προσαρμόζεται για να χρησιμοποιηθεί με τον πιο αποτελεσματικό τρόπο στη φάση της κατάτμησης. Οι τεχνικές προεπεξεργασίας μπορεί να είναι διαφορετικές και να είναι είτε κλιμάκωση (scaling), είτε μετάφραση (translation) είτε υποδειγματοληψία (subsampling). Παρ' όλα αυτά η πρώτη λειτουργία που πρέπει να πραγματοποιηθεί είναι η εξομάλυνση διότι καθιστά δυνατή την εκτέλεση των διαφόρων σκηνών του συνόλου δεδομένων όσο το δυνατόν πιο παρόμοια μεταξύ τους. Οι μελετητές θέλησαν να κάνουν τις σκηνές όσο το δυνατόν πιο παρόμοιες με τις σκηνές δεδομένων ScanNET, που χρησιμοποιούνται από τους συντάκτες του δικτύου PointNet ++ για να τις εκπαιδεύσουν στη φάση της σημασιολογικής κατάτμησης.

Σχετικά με την τεχνική της μετάφρασης, δεδομένου ότι οι συντεταγμένες της σκηνής είναι γεωγραφικές, υπάρχει ο κίνδυνος μερικές σκηνές να έχουν μεγάλες συντεταγμένες και άλλες σκηνές μικρές συντεταγμένες. Γι' αυτό οι σκηνές μετατράπηκαν χωρικά στο σημείο (0,0,0).

Στη διαδικασία απόδοσης κλίμακας, είναι απαραίτητο οι σκηνές να λάβουν σωστή κλίμακα έτσι ώστε οι διαστάσεις τους να είναι χωρικά παρόμοιες. Αυτό συμβαίνει διότι το σύνολο δεδομένων Scannet έχει όλες τις σκηνές στο ίδιο επίπεδο, ενώ οι σκηνές στην μελέτη είναι μνημεία τοποθετημένα σε διαφορετικά επίπεδα. Γι' αυτό, πρέπει να μετατραπούν με τις κατάλληλες αναλογίες ώστε να σέβονται τις διαστάσεις του πλαισίου Scannet.

Σχετικά με την υποδειγματοληψία, τα αρχικά νέφη σημείων αποτελούνται γενικά από εκατομμύρια σημεία. Αυτά είναι πολύ βαριά για να υπολογιστούν από ένα δίκτυο και σίγουρα θα είναι επίσης πολύ περιττά, ακόμη και λαμβάνοντας υπόψη τα γεωμετρικά χαρακτηριστικά ενός αντικειμένου. Έτσι πραγματοποιήθηκε μια δειγματοληψία των νεφών σημείων, μειώνοντας ακόμη και τα αρχικά δεδομένα 100 φορές. Μια περαιτέρω λειτουργία προεπεξεργασίας που θα μπορούσε να βελτιώσει την απόδοση είναι η δυνατότητα εξισορρόπησης των κατηγοριών δεδομένων. Η εξισορρόπηση θα μπορούσε να γίνει από σημείο, προσπαθώντας να εξισορροπήσετε όλα τα σημεία των τάξεων και από αντικείμενο, εξισορροπώντας τον αριθμό των αντικειμένων για κάθε κλάση. Η εξισορρόπηση μπορεί να γίνει με την περαιτέρω υποδειγματοληψία των πυκνών αντικειμένων ή την εξάλειψη αυτών των περιττών αντικειμένων.

3.1.1.3 Σημασιολογική Κατάτμηση PointNet++

Το PointNet++ είναι ένα ιεραρχικό νευρωνικό δίκτυο που επεξεργάζεται ένα σύνολο σημείων δειγματοληπτικά σε έναν μετρικό χώρο. Η βασική ιδέα είναι η εξής:

- Διαχωρίζει το σύνολο των σημείων το οποίο βρίσκεται σε επικαλυπτόμενες τοπικές περιοχές με την απόσταση που έχει από την κάθε μια,
- Παρόμοια με τα συνελκτικά νευρωνικά δίκτυα (CNN), εξάγει τοπικά χαρακτηριστικά αποτυπώνοντας γεωμετρικές δομές από μικρές γειτονιές,
- Αυτά τα τοπικά χαρακτηριστικά ομαδοποιούνται περαιτέρω σε μεγαλύτερες μονάδες και υποβάλλονται σε επεξεργασία για την παραγωγή χαρακτηριστικών υψηλότερου επιπέδου.

Η διαδικασία αυτή επαναλαμβάνεται μέχρι να ληφθούν τα χαρακτηριστικά ολόκληρου του συνόλου των σημείων. Ωστόσο, τα σύνολα σημείων παίρνονται δειγματοληπτικά συνήθως με διαφορετικές πυκνότητες, γεγονός που έχει ως αποτέλεσμα σημαντικά μειωμένη απόδοση για τα δίκτυα που εκπαιδεύονται σε ομοιόμορφες πυκνότητες. Για να αντιμετωπιστεί αυτό το ζήτημα, χρησιμοποιείται ένα σύνολο επιπέδων μάθησης για τον προσαρμοστικό συνδυασμό χαρακτηριστικών από πολλαπλές κλίμακες, δημιουργώντας μια ιεραρχική δομή. Η δομή αυτή αποτελείται από έναν αριθμό επιπέδων αφαίρεσης συνόλων. Σε κάθε επίπεδο, ένα σύνολο σημείων υπόκειται σε επεξεργασία και αφαιρείται για να παραχθεί ένα νέο σύνολο με λιγότερα στοιχεία. Το δίκτυο μπορεί να λειτουργήσει σε 3 πιθανές διαμορφώσεις:

- Κάθε επίπεδο αφαίρεσης περιέχει ομαδοποίηση και εξαγωγή χαρακτηριστικών μιας ενιαίας κλίμακας (Single Scale Grouping).
- Κάθε επίπεδο αφαίρεσης εξάγει πολλαπλές κλίμακες τοπικών μοτίβων και τις συνδυάζει έξυπνα σύμφωνα με τις πυκνότητες των τοπικών σημείων. Όσον αφορά στην ομαδοποίηση τοπικών περιοχών και τον συνδυασμό χαρακτηριστικών από διαφορετικές κλίμακες, μπορούμε να επιλέξουμε δύο τύπους προσαρμοστικών στρωμάτων πυκνότητας, όπως αναφέρονται παρακάτω:
 - Ομαδοποίηση πολλαπλών κλιμάκων (Multi-Scale Grouping),
 - Ανάλυση πολλαπλών κλιμάκων (Multi-Scale Resolution).

Το δίκτυο PointNet ++ έχει 3 διαφορετικές υλοποιήσεις για την πραγματοποίηση 3 διαφορετικών εργασιών: ταξινόμηση, διαχωρισμό μέρους και σημασιολογική κατάτμηση. Στη μελέτη που παρουσιάζεται χρησιμοποιήθηκε ο τρίτος τύπος δικτύου.

Στην αρχή επιλέχθηκαν ποιες σκηνές θα είναι μέρος των δοκιμών. Γενικά, πρέπει να χωριστεί το σύνολο δεδομένων σε εκπαίδευση και δοκιμές, ενδεχομένως να υπάρχουν και οι δύο κατηγορίες αντικειμένων. Η βέλτιστη κατάσταση θα ήταν να έχουν ισορροπημένες τάξεις, αλλά αυτό δεν είναι πάντα δυνατό. Στη συγκεκριμένη μελέτη ωστόσο, θεωρούνται 3 δείγματα σκηνών: από κοινού έχουν μόνο το παράθυρο, το τοίχο, το δάπεδο, τη στέγη. Στην πραγματικότητα, στην περίπτωση που οι σκηνές έχουν διαφορετικούς τύπους αντικειμένων, οι τάξεις τέμνονται, αφήνοντας μόνο τα αντικείμενα που είναι κοινά σε όλους. Σε αυτή την περίπτωση είναι χρήσιμη η υποδιαίρεση των αντικειμένων στους διάφορους φακέλους, καθώς με αυτόν τον τρόπο χάρη σε ένα σενάριο Python, οι σκηνές μπορούν να δημιουργηθούν με εξατομικευμένο τρόπο (Malinverni κ.ά. 2019).

Έπειτα ο κώδικας Scannet τροποποιείται για τη συγκεκριμένη περίπτωση:

- Οι σκηνές εκπαίδευσης και δοκιμών μπορούν να εισαχθούν ως λίστες των ονομάτων.
- Εκτός από τις συντεταγμένες XYZ, μπορούν επίσης να διαβαστούν τα άλλα χαρακτηριστικά (χρώμα, ανακλαστικότητα) από τα αρχεία σκηνής, απλά ρυθμίζοντας τις αρχικές παραμέτρους.
- Η αρχική έκδοση του PointNet ++ έκανε μια εκπαίδευση στην οποία ένας δειγματολήπτης πήρε μόνο 8192 πόντους για κάθε σκηνή εκπαίδευσης. Αλλά με αυτόν τον τρόπο η εκπαίδευση ήταν ανισόρροπη, γιατί το δίκτυο ήταν μη ισορροπημένο μόνο στις τάξεις που ελήφθησαν από αυτόν τον δειγματολήπτη. Αντ' αυτού στη παρούσα μελέτη δοκιμάζεται ολόκληρη η σκηνή εκπαίδευσης (όπως συμβαίνει και για τη δοκιμή).
- Κατά τη διάρκεια της εκπαίδευσης αποθηκεύονται:
 - τα βάρη του δικτύου.
 - Ο πίνακας σύγχυσης και οι μετρήσεις (ακρίβεια, επαναφορά, F1-score) της εκπαίδευσης και των δοκιμών.
 - Οι μετρήσεις που βασίζονται σε σημεία και μετρήσεις που βασίζονται σε ογκομετρικά στοιχεία (voxels)
 - τα δεδομένα εισόδου στο δίκτυο και τις εξόδους που προβλέπεται από το δίκτυο (δηλ. νέφη σημείων με τις προαναφερθείσες ετικέτες).

Η περιγραφόμενη μέθοδος ταξινόμησης βασίζεται σε σημεία, δηλ. υπάρχει μια προβλεπόμενη τάξη για κάθε σημείο. Για να αξιολογηθεί περαιτέρω η προσέγγιση οι μελετητές πραγματοποίησαν μια φάση μετά την επεξεργασία για να μετατρέψουν τα προβλεπόμενα αποτελέσματα σε μια ταξινόμηση αντικειμένων. Χρησιμοποίησαν μια προσέγγιση Winner Take All (WTA). Ως πρώτο βήμα, απομόνωσαν τα μεμονωμένα αντικείμενα από τη δοκιμαστική σκηνή χρησιμοποιώντας τα πλαίσια οριοθέτησης που είχαν προηγουμένως υπολογιστεί. Στη συνέχεια, υπολόγισαν το ποσοστό ταξινόμησης κάθε προβλεπόμενης κλάσης για κάθε μεμονωμένο αντικείμενο. Εάν το ποσοστό της ταξινόμησης υπερβαίνει ένα συγκεκριμένο όριο, τότε σχετίζεται αυτή η τάξη με αυτό το αντικείμενο.

3.1.1.4 Αξιολόγηση αποτελεσμάτων

Χάρη σε όλα αυτά που αποθηκεύονται κατά τη διάρκεια της εκπαίδευσης, είναι δυνατόν να υπάρχει η γνώση του αν το δίκτυο είναι καλά εκπαιδευμένο ή η μη ισορροπία που εμφανίζεται αφορά μόνο σε μερικές τάξεις. Με αυτόν τον τρόπο οι παράμετροι PointNet ++ μπορούν να τροποποιηθούν προκειμένου να βελτιωθούν τα αποτελέσματα. Πραγματικά λαμβάνοντας υπόψη τη ροή εργασίας της εικόνας 10 (βλ. σελ 24), θα είναι επίσης δυνατό η επιστροφή στα προηγούμενα βήματα και να πραγματοποιηθούν αλλαγές στην προεπεξεργασία του συνόλου δεδομένων ή ακόμα και στην επισήμανση. Έτσι, η ροή εργασίας μπορεί να θεωρηθεί ως κύκλος (με ανατροφοδότηση), όχι απαραίτητα ως διαδοχική διαδικασία (Malinverni κ.ά. 2019).

3.2 Αστικού Τύπου Δεδομένα

Υπάρχουν πολλές μελέτες όπου εφαρμόζονται μέθοδοι σημασιολογικής κατάτμησης σε αστικού τύπου δεδομένα. Στην παρούσα εργασία παρουσιάζεται η μελέτη των Adam κ.ά. (2019) διότι εφαρμόζουν μια υβριδική μέθοδο η οποία βασίζεται τόσο στον δισδιάστατο όσο και στον τρισδιάστατο χώρο και αποσκοπεί στη διερεύνηση του κατά πόσον η επιλογή εικόνας είναι κρίσιμη όσον αφορά στην ακρίβεια της τρισδιάστατης σημασιολογικής κατάτμησης νέφους σημείων. Τα πειραματικά αποτελέσματα παρουσιάζονται σε ένα δωρεάν διαδικτυακό σύνολο δεδομένων που απεικονίζει οικοδομικά τετράγωνα γύρω από το Παρίσι. Η πειραματική διαδικασία όχι μόνο επικυρώνει ότι τα υβριδικά χαρακτηριστικά (γεωμετρικά και οπτικά) μπορούν να επιτύχουν ακριβέστερη σημασιολογική κατάτμηση, αλλά και καταδεικνύει τη σημασία της καταλληλότερης προβολής για την εξαγωγή 2D χαρακτηριστικών (Adam κ.ά. 2019).

Έχουν υπάρξει και άλλες μέθοδοι που συνδυάζουν πληροφορίες από τον δισδιάστατο και τον τρισδιάστατο χώρο με στόχο τη σημασιολογική κατάτμηση. Αυτές οι μέθοδοι βασίζονται σε προσεγγίσεις που βασίζονται στη μάθηση και απαιτούν επισημασμένα δεδομένα εκπαίδευσης. Οι Landrieu κ.ά. το 2017 πραγματοποιούν μάθηση συνόλου, με βάση τρισδιάστατα χαρακτηριστικά από το νέφος σημείων και δισδιάστατα χαρακτηριστικά που προέρχονται από την προβολή των σημείων σε ένα οριζόντιο προσανατολισμένο επίπεδο. Οι Ladicky κ.ά. το 2012, στοχεύουν επίσης στη σημασιολογική κατάτμηση τρισδιάστατων σκηνών, συνδυάζοντας χάρτες βάθους και χαρακτηριστικά εμφάνισης. Οι Serna και Marcotegui το 2014, χρησιμοποίησαν εικόνες υψομέτρων και τα τελικά αποτελέσματα προκύπτουν από την επαναπροβολή των εικόνων στο τρισδιάστατο νέφος σημείων. Τα αντικείμενα ταξινομούνται με τη βοήθεια μηχανών διανυσμάτων υποστήριξης (Support Vector Machines).

Επιπλέον, υπάρχουν και κάποιες εργασίες, που προσπαθούν να επιτύχουν μια ακριβή κατάτμηση μοντέλων πόλεων και πιο συγκεκριμένα αφορούν στην ανάλυση προσόψεων. Αυτές οι προσεγγίσεις αποσκοπούν στην κατάτμηση των διαφόρων τμημάτων μιας πρόσοψης και στην απόδοση σε αυτά μιας σημασιολογικής ετικέτας, όπως πόρτα, παράθυρο, μπαλκόνι κ.λ.π. με τη βοήθεια σχολιασμένων δεδομένων. Μια μεθοδολογία η οποία σχετίζεται αρκετά με αυτή που θα παρουσιαστεί σχηματίζει ένα υβριδικό διάνυσμα, το οποίο περιλαμβάνει τόσο δισδιάστατα όσο και τρισδιάστατα χαρακτηριστικά που εισάγονται στο σχήμα ταξινόμησης. Τα δισδιάστατα χαρακτηριστικά (τιμές RGB και Lab) εξάγονται ως οι μέσες τιμές μεταξύ όλων των εικόνων στις οποίες είναι ορατό το προς ταξινόμηση σημείο, ενώ τα τρισδιάστατα χαρακτηριστικά όπως το ύψος και το βάθος του σημείου καθώς και οι διανυσματικές κανονικές εξάγονται απευθείας από το πλέγμα. Τέλος, σε παρόμοια κατεύθυνση με την μελέτη που θα παρουσιαστεί είναι και μια άλλη έρευνα η οποία χρησιμοποιεί πλέγματα αντί για νέφη σημείων και επιλέγει την καλύτερη προβολή για κάθε τριγωνικό πλέγμα προς ταξινόμηση και χρησιμοποιεί αυτή την εικόνα για την εξαγωγή 2D χαρακτηριστικών.

3.2.1 Μεθοδολογία

Το πρόβλημα που καλείται να αντιμετωπίσει η μελέτη που θα παρουσιαστεί είναι ένα κλασικό πρόβλημα ταξινόμησης, δηλαδή από ένα σύνολο δεδομένων να εκτιμηθεί σωστά η κατηγορία στην οποία ανήκει το κάθε αντικείμενο. Η συνεισφορά της μελέτης αυτής είναι η διατύπωση ενός ευρετικού κανόνα για την βέλτιστη αναγνώριση εικόνων (Adam κ.ά. 2019).

3.2.1.1 Δεδομένα

Στην μελέτη (Adam κ.ά. 2019), χρησιμοποιήθηκε ένα ελεύθερα διαθέσιμο σύνολο δεδομένων, το οποίο αποτελείται από 428 εικόνες που αποτυπώνουν αστικές σκηνές. Τα δεδομένα δημιουργήθηκαν με τις τεχνικές Structure for Motion (SfM) και Multiview Stereo (MvS) επομένως τα δεδομένα εικόνας είναι διαθέσιμα μαζί με το νέφος σημείων. Επειδή η μεθοδολογία που προτείνεται στην μελέτη στοχεύει στη σημασιολογική κατάτμηση νεφών σημείων, ενσωματώνοντας τόσο δισδιάστατα όσο και τρισδιάστατα χαρακτηριστικά, χρησιμοποιούνται μόνο οι εικόνες και τα νέφη σημείων. Το σύνολο δεδομένων περιέχει οκτώ διαφορετικές σημασιολογικές κλάσεις, οι οποίες είναι ουρανός, μπαλκόνι, πρόσοψη, κατάσταση, παράθυρο, πόρτα, στέγη και μια κενή κλάση, μαζί με ένα διαχωρισμό εκπαίδευσης/δοκιμής και ένα πρωτόκολλο αξιολόγησης. Ένα παράδειγμα του χρησιμοποιούμενου συνόλου δεδομένων απεικονίζεται στην Εικόνα 13, όπου απεικονίζεται ο διαχωρισμός των διαφόρων σημασιολογικών κλάσεων. Πιο συγκεκριμένα, η τάξη κατάσταση απεικονίζεται με πράσινο χρώμα, το μπαλκόνι με ματζέντα, ο ουρανός με γαλάζιο, η στέγη με σκούρο μπλε, το παράθυρο με κόκκινο, η πόρτα με πορτοκαλί και η κύρια πρόσοψη με κίτρινο. Τα σημεία που ανήκουν στην κενή κατηγορία απεικονίζονται με μαύρο χρώμα.



Εικόνα 13: Διαχωρισμός διαφόρων σημασιολογικών κλάσεων με χρώματα

3.2.1.2 Εξαγωγή χαρακτηριστικών

Η καινοτομία αυτής της μελέτης είναι ότι επικεντρώνεται στην πρόταση μιας καινοτόμου τεχνικής βέλτιστης επιλογής προβολής για την εξαγωγή δισδιάστατων χαρακτηριστικών, όταν πρόκειται να ταξινομηθούν νέφη σημείων αντί για πλέγματα. Όταν επιλέγεται η βέλτιστη προβολή, δηλαδή αυτή που έχει τις περισσότερες πιθανότητες να ταξινομήσει σωστά το επιθυμητό αντικείμενο, η συνολική ακρίβεια της σημασιολογικής κατάτμησης βελτιώνεται.

Για τη βέλτιστη επιλογή προβολής, πρώτα όλα τα τρισδιάστατα σημεία X προβάλλονται, μέσω των πινάκων προβολής P στο δισδιάστατο χώρο, σύμφωνα με την παρακάτω εξίσωση και επιλέγονται οι εικόνες στις οποίες τα σημεία είναι ορατά:

$$X = P \cdot x \quad (1)$$

όπου το X αναφέρεται στις τρισδιάστατες συντεταγμένες του σημείου και το x στο σημείο της εικόνας 2D. Τέλος, το P αναφέρεται στον πίνακα προβολής.

Έπειτα, γύρω από κάθε σημείο στον τρισδιάστατο χώρο σχηματίζεται ένα ορθογώνιο τμήμα στο επίπεδο X , με προκαθορισμένες διαστάσεις. Οι συντεταγμένες των τεσσάρων γωνιών που σχηματίζουν το ορθογώνιο τμήμα είναι γνωστές και προβάλλονται μέσω της εξίσωσης 1 στις διάφορες εικόνες. Από όλες τις εικόνες στις οποίες το σημείο είναι ορατό, επιλέγονται μόνο εκείνες στις οποίες απεικονίζεται ολόκληρο το ορθογώνιο (δηλαδή και οι τέσσερις γωνίες του προβάλλονται στην εικόνα). Στη συνέχεια, υπολογίζεται το εμβαδόν του προβαλλόμενου τμήματος σε κάθε εικόνα και η βέλτιστη εικόνα είναι αυτή που εξασφαλίζει το μέγιστο εμβαδόν του ορθογωνίου (δηλαδή η εικόνα με τον μέγιστο αριθμό εικονοστοιχείων που σχηματίζουν το προβαλλόμενο τμήμα). Χρησιμοποιώντας αυτόν τον ευρετικό κανόνα για τη βέλτιστη επιλογή εικόνας, αποκλείονται οι πολύ λοξές και μικρής κλίμακας προβολές για το βήμα εξαγωγής χαρακτηριστικών 2D.

Όταν εντοπιστεί η βέλτιστη προβολή, εξάγονται χαρακτηριστικά. Όσον αφορά στα γεωμετρικά (τρειςδιάστατα) χαρακτηριστικά, το βάθος (d), το κανονικοποιημένο ύψος (h) και το κάθετο διάνυσμα (n) για κάθε σημείο υπολογίζονται απευθείας από το νέφος σημείων. Από την άλλη πλευρά, τα χαρακτηριστικά υφής εξάγονται από τη βέλτιστη εικόνα και περιλαμβάνουν τις τιμές RGB και Lab του προβαλλόμενου σημείου. Ένα υβριδικό διάνυσμα:

$$V = d \quad h \quad n \quad RGB \quad Lab \quad (2)$$

με διαστάσεις 11x1 σχηματίζεται για κάθε σημείο και στη συνέχεια εισάγεται σε έναν ταξινομητή Random Forest (RF) και έναν ταξινομητή AdaBoost.

3.2.1.3 Εφαρμογή τεχνικών μηχανικής μάθησης

Προκειμένου να αντιμετωπιστεί η πολυπλοκότητα του χώρου των χαρακτηριστικών εισόδου, χρησιμοποιείται ο ταξινομητής τυχαίου δάσους (Random Forest). Ο ταξινομητής τυχαίου δάσους είναι μια μέθοδος συνόλου δέντρων απόφασης, η οποία έχει υψηλή ακρίβεια σε διάφορους τομείς. Τα δέντρα αποφάσεων, ως μοντέλα πρόβλεψης, αντιστοιχίζουν τις παρατηρήσεις σχετικά με ένα στοιχείο σε συμπεράσματα σχετικά με την τιμή-στόχο του στοιχείου. Στις δενδρικές δομές ταξινόμησης, τα "φύλλα" αντιπροσωπεύουν τις ετικέτες κλάσεων και οι κλάδοι αντιπροσωπεύουν τις συνδέσεις των χαρακτηριστικών που οδηγούν σε αυτές τις ετικέτες κλάσεων. Κάθε εσωτερικός κόμβος επισημαίνεται με ένα χαρακτηριστικό εισόδου. Τα τόξα που προέρχονται από έναν κόμβο που φέρει ετικέτα με ένα χαρακτηριστικό, φέρουν ετικέτα με κάθε μία από τις πιθανές τιμές του χαρακτηριστικού. Κάθε φύλλο του δέντρου επισημαίνεται με μια κλάση ή μια κατανομή πιθανότητας πάνω στις κλάσεις. Οι αλγόριθμοι για την κατασκευή δέντρων αποφάσεων λειτουργούν συνήθως από πάνω προς τα κάτω, επιλέγοντας σε κάθε βήμα μια μεταβλητή που χωρίζει καλύτερα το σύνολο των στοιχείων. Διαφορετικοί αλγόριθμοι χρησιμοποιούν διαφορετικές μετρικές για τη μέτρηση του "καλύτερου". Αυτές γενικά μετρούν την ομοιογένεια της μεταβλητής-στόχου εντός των υποσυνόλων. Αυτές οι μετρικές εφαρμόζονται σε κάθε υποψήφιο υποσύνολο και οι τιμές που προκύπτουν συνδυάζονται (π. χ. υπολογίζονται κατά μέσο όρο) για να παρέχουν ένα μέτρο της ποιότητας του διαχωρισμού (Adam κ.ά. 2019).

Το Adaptive boosting (AdaBoost) είναι ένας αλγόριθμος μάθησης συνόλου, ο οποίος είναι πιο ανθεκτικός στην υπερπροσαρμογή, αλλά είναι συχνά ευαίσθητος σε θορυβώδη δεδομένα και ακραίες τιμές. Το AdaBoost δημιουργεί έναν ισχυρό μαθητή (έναν ταξινομητή που συσχετίζεται καλά με τον πραγματικό ταξινομητή) προσθέτοντας επαναληπτικά αδύναμους μαθητές (έναν ταξινομητή που συσχετίζεται ελάχιστα με τον πραγματικό ταξινομητή). Κατά τη διάρκεια κάθε γύρου εκπαίδευσης, ένας νέος αδύναμος μαθητής προστίθεται στο σύνολο και ένα διάνυμα στάθμισης προσαρμόζεται ώστε να εστιάζει σε παραδείγματα που είχαν ταξινομηθεί λανθασμένα σε προηγούμενους γύρους. Το αποτέλεσμα είναι ένας ταξινομητής που έχει μεγαλύτερη ακρίβεια από τους ταξινομητές των αδύναμων μαθητών (Adam κ.ά. 2019).

3.2.1.4 Αποτελέσματα

Τα αποτελέσματα της μελέτης παρουσιάστηκαν ποσοτικά για την αξιολόγηση της αποτελεσματικότητας της προτεινόμενης μεθοδολογίας. Η σημασιολογική κατάτμηση εφαρμόστηκε ξεχωριστά στο δισδιάστατο και τρισδιάστατο χώρο σχηματίζοντας ένα διάνυμα αποκλειστικά από τα προαναφερθέντα τρισδιάστατα και δισδιάστατα χαρακτηριστικά, αντίστοιχα, αλλά και στο συνδυασμό αυτών των χώρων, όπου ένα υβριδικό διάνυμα που περιέχει τα προαναφερθέντα τρισδιάστατα και δισδιάστατα χαρακτηριστικά (τα δισδιάστατα χαρακτηριστικά εξάγονται από μια τυχαία εικόνα). Για την αξιολόγηση χρησιμοποιήθηκε η μετρική της μέσης ακρίβειας. Παρουσιάζονται πειραματικά αποτελέσματα για τις δύο τεχνικές μηχανικής μάθησης που χρησιμοποιήθηκαν, δηλαδή τις τεχνικές Random Forest και Adaboost. Τα αποτελέσματα παρατίθενται στον πίνακα 1

	AdaBoost (random image)	RF(random image)	AdaBoost (optimal image)	RF (optimal image)
2D features	39.3%	48,6%	46.5%	57.2%
3D features	40.2%	49.9%	40.2%	49.9%
2D+3D features	47.2%	50.1%	54.9%	60.0%

Πίνακας 1: Ακρίβεια ταξινόμησης όταν χρησιμοποιούνται μόνο 2D, μόνο 3D και συνδυασμένα 2D + 3D χαρακτηριστικά για τη βασική και τη βέλτιστη περίπτωση εικόνας, τόσο για τους ταξινομητές Random Forest όσο και για τους ταξινομητές AdaBoost.

Όπως φαίνεται και από τον πίνακα ο συνδυασμός των 2D και 3D χαρακτηριστικών οδήγησε σε μια πιο ακριβή σημασιολογική κατάτμηση της 3D σκηνής. Επιπλέον, επιβεβαιώθηκε η αρχική υπόθεση της μελέτης, ότι δηλαδή μια πιο εξελιγμένη διαδικασία επιλογής εικόνων θα οδηγούσε σε πιο ακριβή αποτελέσματα σημασιολογικής κατάτμησης. Ειδικότερα, η συγχώνευση των 2D και 3D χαρακτηριστικών οδήγησε σε μια συνολική μέση ακρίβεια ίση με 50,1% για τη βασική διαδικασία και μια μέση ακρίβεια ίση με 60,0% για την προτεινόμενη διαδικασία, όταν χρησιμοποιούνται Random Forests.

Επιπλέον, μετά από μια προσεκτική ματιά στα ποσοτικά αποτελέσματα, μπορεί κανείς να παρατηρήσει ότι η επιλογή της εικόνας είναι ένα εξαιρετικά κρίσιμο βήμα προς τη σημασιολογική κατάτμηση της σκηνής, δεδομένου ότι στις περιπτώσεις που χρησιμοποιούνται ξεχωριστά 2D και 3D χαρακτηριστικά. Πιο συγκεκριμένα, μόνο τα 2D χαρακτηριστικά οδήγησαν σε σχετικά υψηλότερα ποσοστά ακρίβειας όταν επιλέχθηκε η βέλτιστη εικόνα, δηλαδή 57,2% ακρίβεια σε σύγκριση με 49,9% όταν χρησιμοποιήθηκαν μόνο 3D χαρακτηριστικά. Από την άλλη πλευρά, όταν επιλέχθηκε μια τυχαία εικόνα για την εξαγωγή 2D χαρακτηριστικών, τα 3D χαρακτηριστικά φαίνεται να είναι πιο σημαντικά, καθώς οδηγούν σε ακρίβεια 49,9%, σε σύγκριση με το 48,6% που επιτυγχάνεται με τα 2D χαρακτηριστικά.

Τέλος, σύμφωνα με την μελέτη (Adam κ.ά. 2019) είναι σημαντικό να σημειωθεί ότι πιο εξελιγμένες τεχνικές μηχανικής μάθησης, όπως τα Conditional Random Fields (CRF) θα μπορούσαν να ενσωματωθούν στην όλη διαδικασία για την επίτευξη υψηλότερων ποσοστών ακρίβειας.

4. ΣΥΜΠΕΡΑΣΜΑΤΑ

Με την εξέλιξη της τεχνολογίας έχει αναπτυχθεί σε μεγάλο βαθμό η φωτογραμμετρία έχοντας ως αποτέλεσμα την δημιουργία νέων μορφών δεδομένων και νέες ανάγκες επεξεργασίας τους. Μια τέτοια μορφή δεδομένων είναι τα νέφη σημείων. Υπάρχουν διάφοροι τρόποι για να δημιουργηθεί ένα νέφος σημείων. Ένας τρόπος είναι απ' ευθείας από σαρωτές (laser scanner). Ένας άλλος είναι έμμεσα μέσω λογισμικού φωτογραμμετρίας. Σε κάθε περίπτωση όμως αποτελούν σήμερα μια μορφή δεδομένων πολύ χρήσιμη ταυτόχρονα όμως και πολύ ελκυστική για τη διερεύνηση μεθόδων και τεχνικών ώστε να τα επεξεργαζόμαστε πιο εύκολα και πιο αυτόματα. Έτσι, με τα χρόνια δημιουργήθηκε η ανάγκη για διερεύνηση μεθόδων για τη σημασιολογική κατάτμηση ενός νέφους σημείων.

Σημασιολογική κατάτμηση ενός νέφους σημείων είναι η διαδικασία κατά την οποία τα σημεία από τα οποία αποτελείται ένα νέφος σημείων χωρίζονται σε κατηγορίες έχοντας κοινά χαρακτηριστικά και μια ιδιότητα. Πριν εφαρμοστούν ευρέως αποτελεσματικές μέθοδοι μάθησης με επίβλεψη στη σημασιολογική κατάτμηση, η κατάτμηση νέφους σημείων χωρίς επίβλεψη ήταν ένα σημαντικό έργο για δεδομένα 2,5D/3D. Η κατάτμηση νέφους σημείων στοχεύει στην ομαδοποίηση σημείων με παρόμοια γεωμετρικά/φασματικά χαρακτηριστικά χωρίς να λαμβάνει υπόψη τις σημασιολογικές πληροφορίες.

Οι τεχνικές κατάτμησης χωρίζονται σε τέσσερις κατηγορίες, με βάση τις ακμές, την ανάπτυξη περιοχών, την προσαρμογή μοντέλων και την ομαδοποίηση. Σύμφωνα με αυτά που αναφέρθηκαν η τεχνική με βάση τις ακμές σπάνια εφαρμόζεται σήμερα για πυκνά και μεγάλα σε έκταση δεδομένα νέφη σημείων καθώς τα αποτελέσματα δεν είναι καθόλου καλά. Είναι όμως μια καλή μέθοδος για γρήγορη κατάτμηση σε νέφη σημείων με απλές σκηνές και ιδανικά σημεία, δηλαδή με χαμηλό θόρυβο και ομοιόμορφη πυκνότητα. Αντίθετα η τεχνική με βάση την ανάπτυξη περιοχών είναι μια κλασική μέθοδος η οποία χρησιμοποιείται ακόμα και σήμερα. Το ίδιο συμβαίνει και με τις άλλες δυο τεχνικές, δηλαδή με βάση την προσαρμογή μοντέλων και την ομαδοποίηση και συνεχώς γίνονται νέες μελέτες βρίσκοντας τρόπους να βελτιώσουν τους αλγορίθμους ώστε να μειωθούν όσο γίνεται περισσότερο τα μειονεκτήματά τους. Κοινό στοιχείο όμως όλων αυτών των τεχνικών είναι ότι τα σημεία που αποτελούν το νέφος σημείων ομαδοποιούνται χωρίς να δίνεται ερμηνεία του τι είναι η κάθε ομάδα. Παρ' όλα αυτά αποτελούν ένα βασικό στάδιο για την εξέλιξη της κατάτμησης σε σημασιολογική κατάτμηση δηλαδή στο να δίνεται και η ιδιότητα των σημείων. Αυτό σε συνδυασμό με την εξέλιξη της τεχνολογίας αναπτύχθηκαν τεχνικές για τη σημασιολογική κατάτμηση των νέφων σημείων.

Οι τεχνικές σημασιολογικής κατάτμησης χωρίστηκαν σε δυο βασικές κατηγορίες, στην κλασική επιβλεπόμενη μάθηση και στη βαθιά μάθηση. Ωστόσο παρατηρείται ένα ιδιαίτερο ενδιαφέρον στην ανάπτυξη μεθόδων με βαθιά μάθηση καθώς αποτελεί μια από τις πιο πρόσφατες και γρήγορα εξελισσόμενες τεχνικές στην αναγνώριση προτύπων, με αποτέλεσμα να παρατηρούνται όλο και καλύτερα αποτελέσματα στη σημασιολογική κατάτμηση.

Έχοντας παρουσιάσει εν συντομία τις μεθόδους κατάτμησης και σημασιολογικής κατάτμησης, εύκολα καταλαβαίνει κανείς ότι η σημασιολογική κατάτμηση ενός νέφους σημείων είναι αρκετά πιο δύσκολη από την δισδιάστατη σημασιολογική κατάτμηση. Επίσης, ενώ υπάρχουν πολλές προσεγγίσεις καθεμία από αυτές έχει δοκιμαστεί σε πολύ περιορισμένα και ανόμοια σύνολα δεδομένων, οπότε είναι δύσκολο να επιλεγεί η βέλτιστη προσέγγιση για πρακτικές εφαρμογές. Οι μέθοδοι που βασίζονται στη βαθιά μάθηση φαίνεται να κερδίζουν ολοένα και περισσότερο έδαφος, ωστόσο μια συνεχής βελτίωση και η δημιουργία ενός προτύπου νευρωνικού δικτύου θα βοηθήσει στο πρόβλημα της σημασιολογικής κατάτμησης.

Τέλος, για καλύτερη σύγκριση της αποτελεσματικότητας των μεθόδων θα μπορούσαν να πραγματοποιηθούν μελέτες οι οποίες θα χρησιμοποιούν το ίδιο νέφος σημείων εφαρμόζοντας τις διάφορες μεθόδους.

5. ΒΙΒΛΙΟΓΡΑΦΙΑ

Ξενόγλωσση

Adam, A., Grammatikopoulos, L., Karras, G., Protopapadakis, E. και Karantzas, K. 2019. A semantic 3D point cloud segmentation approach based on optimal view selection for 2D image feature extraction. Στο: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. International Society for Photogrammetry and Remote Sensing, σσ. 9–14. doi: 10.5194/isprs-archives-XLII-2-W17-9-2019.

Boulch, A., Guerry, J., le Saux, B. και Audebert, N. 2018. SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks. *Computers and Graphics (Pergamon)* 71, σσ. 189–198. doi: 10.1016/j.cag.2017.11.010.

Castillo, E., Liang, J. και Zhao, H. 2013. Point Cloud Segmentation and Denoising via Constrained Nonlinear Least Squares Normal Estimates. Στο: *Innovations for Shape Analysis.*, σσ. 283–299. doi: 10.1007/978-3-642-34141-0_13.

Choi, S., Kim, T. και Yu, W. 2009. Performance evaluation of RANSAC family. Στο: *British Machine Vision Conference, BMVC 2009 - Proceedings*. British Machine Vision Association, BMVA. doi: 10.5244/C.23.81.

Deschaud, J.-E. και Goulette, F. 2010. *A Fast and Accurate Plane Detection Algorithm for Large Noisy Point Clouds Using Filtered Normals and Voxel Growing*. Διαθέσιμο στο: <https://hal-mines-paristech.archives-ouvertes.fr/hal-01097361>.

Dong, Z., Yang, B., Hu, P. και Scherer, S. 2018. An efficient global energy optimization approach for robust 3D plane segmentation of point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 137, σσ. 112–133. doi: 10.1016/j.isprsjprs.2018.01.013.

Duda, R.O. και Hart, P.E. 1972. *Graphics and Use of the Hough Transformation To Detect Lines and Curves in Pictures*.

Grilli, E., Menna, F. και Remondino, F. 2017. A review of point clouds segmentation and classification algorithms. Στο: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. International Society for Photogrammetry and Remote Sensing, σσ. 339–344. doi: 10.5194/isprs-archives-XLII-2-W3-339-2017.

Landrieu, L., Raguét Raguét, H., Vallet, B., Mallet, C., Weinmann, M., Landrieu, L. και Raguét, H. 2017. A structured regularization framework for spatially smoothing semantic labelings of 3D point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 132, σσ. 102–118. Διαθέσιμο στο: <https://hal.science/hal-01505245v2>.

Landrieu, L. και Simonovsky, M. 2017. Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. Διαθέσιμο στο: <http://arxiv.org/abs/1711.09869>.

Li, C.R.Q., Hao, Y., Leonidas, S. και Guibas, J. 2017. *PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space*.

Lyne Tchampi, Christopher B Choy, Iro Armeni, JunYoung Gwak, Silvio Savarese, 2017. *SEGCloud: Semantic Segmentation of 3D Point Clouds*. Διαθέσιμο στο: <http://segcloud.stanford.edu>.

Malinverni, E.S., Pierdicca, R., Paolanti, M., Martini, M., Morbidoni, C., Matrone, F. και Lingua, A. 2019. DEEP LEARNING for SEMANTIC SEGMENTATION of 3D POINT CLOUD. Στο: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. International Society for Photogrammetry and Remote Sensing, σσ. 735–742. doi: 10.5194/isprs-archives-XLII-2-W15-735-2019.

Maturana, D. και Scherer, 2015. *VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition*.

- Nguyen, A. και Le, B. 2013. 3D point cloud segmentation: A survey. Στο: *IEEE Conference on Robotics, Automation and Mechatronics, RAM - Proceedings*. IEEE Computer Society, σσ. 225–230. doi: 10.1109/RAM.2013.6758588.
- Ning, X., Zhang, X., Wang, Y. και Jaeger, M. 2009. Segmentation of architecture shape information from 3D point cloud. Στο: *Proceedings - VRCAI 2009: 8th International Conference on Virtual Reality Continuum and its Applications in Industry.*, σσ. 127–132. doi: 10.1145/1670252.1670280.
- Pfeifer, N., Tóvári, D. και Pfeifer, N. 2012. *Segmentation based robust interpolation-A new Approach to laser data filtering* Διαθέσιμο στο: <https://www.researchgate.net/publication/228927546>.
- Qi, C.R., Su, H., Mo, K. και Guibas, L.J. 2016. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Διαθέσιμο στο: <http://arxiv.org/abs/1612.00593>.
- Rabbani, T., Heuvel, F.A. και Vosselman, G. 2006. Segmentation of point clouds using smoothness constraint. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36, Διαθέσιμο στο: <https://www.researchgate.net/publication/228340970>.
- Raguram, R., Frahm, J.M. και Pollefeys, M. 2008. A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus. Στο: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag, σσ. 500–513. doi: 10.1007/978-3-540-88688-4_37.
- Su, H., Maji, S., Kalogerakis, E. και Learned-Miller, E. [χωρίς χρονολογία]. *Multi-view Convolutional Neural Networks for 3D Shape Recognition*. Διαθέσιμο στο: <http://vis-www.cs.umass.edu/mvcnn>.
- Vo, A.V., Truong-Hong, L., Laefer, D.F. και Bertolotto, M. 2015. Octree-based region growing for point cloud segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing* 104, σσ. 88–100. doi: 10.1016/j.isprsjprs.2015.01.011.
- Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M. και Solomon, J.M. 2019. Dynamic graph Cnn for learning on point clouds. *ACM Transactions on Graphics* 38(5). doi: 10.1145/3326362.
- Wang, Z. κ.ά. 2015. A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification. *IEEE Transactions on Geoscience and Remote Sensing* 53(5), σσ. 2409–2425. doi: 10.1109/TGRS.2014.2359951.
- Weinmann, M., Jutzi, B., Hinz, S. και Mallet, C. 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing* 105. doi: 10.1016/j.isprsjprs.2015.01.016.
- Xiao, J., Zhang, J., Adler, B., Zhang, H. και Zhang, J. 2013. Three-dimensional point cloud plane segmentation in both structured and unstructured environments. *Robotics and Autonomous Systems* 61, σσ. 1641–1652. doi: 10.1016/j.robot.2013.07.001.
- Xie, Y., Tian, J. και Zhu, X.X. 2020. Linking Points with Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geoscience and Remote Sensing Magazine* 8(4), σσ. 38–59. doi: 10.1109/MGRS.2019.2937630.
- Zhang, J., Lin, X. και Ning, X. 2013. SVM-Based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sensing* 5(8), σσ. 3749–3775. doi: 10.3390/rs5083749.
- Zhang, J., Zhao, X., Chen, Z. και Lu, Z. 2019. A Review of Deep Learning-Based Semantic Segmentation for Point Cloud. *IEEE Access* 7, σσ. 179118–179133. doi: 10.1109/ACCESS.2019.2958671.

Ελληνική

Πατιάς, Π. (1991). *Εισαγωγή στην Φωτογραμμετρία*. Θεσσαλονίκη, Εκδόσεις "ΖΗΤΗ"

